

BUNDESREPUBLIK DEUTSCHLAND



REC'D	24 NOV 1997
WIPO	PCT

PRIORITY DOCUMENT

Bescheinigung

Herr Professor Dr. Hermann B u j a r d in Heidelberg/
Deutschland hat eine Patentanmeldung unter der Bezeichnung

"Rekombinantes Herstellungsverfahren für ein
vollständiges Malaria-Antigen gp190/MSP1"

am 2. Oktober 1996 beim Deutschen Patentamt eingereicht.

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

Die Anmeldung hat im Deutschen Patentamt vorläufig die Symbole C 12 N, A 61 K und C 07 K der Internationalen Patentklassifikation erhalten.

München, den 13. Oktober 1997

Der Präsident des Deutschen Patentamts

Im Auftrag

Agurks

Aktenzeichen: 196 40 817.2

GRÜNECKER, KINKELDEY, STOCKMAIR & SCHWANHÄUSSER

ANWALTSSOZIETÄT

ANWALTSSOZIETÄT MAXIMILIANSTRASSE 58 D-80538 MÜNCHEN GERMANY

RECHTSANWÄLTE

PATENTANWÄLTE
EUROPEAN PATENT ATTORNEYS

DR. HERMANN SCHWANHÄUSSER
DR. HELMUT EICHMANN
GERHARD BARTH
DR. ULRICH BLUMENRÖDER, LL. M.
CHRISTA NIKLAS-FALTER

AUGUST GRÜNECKER
DR. HERMANN KINKELDEY
DR. WILFRIED STOCKMAIR, AE. E.
DR. KLAUS SCHUMANN
PETER H. JAKOB
DR. GUNTER BEZOLD
WOLFHARD MEISTER
HANS HILGERS
DR. HENNING MEYER-PLATH
ANNELEI EHNOLD
THOMAS SCHUSTER
DR. WALTER LANGHOFF
DR. KLARA GOLDBACH
MARTIN AUFENANGER
GOTTFRIED KLITZSCH
DR. HEIKE VOGELSANG-WENKE
REINHARD KNAUER
DIETMAR KUHL
DR. FRANZ-JOSEF ZIMMER
BETTINA K. REICHELT

IHR ZEICHEN / YOUR REF.

UNSER ZEICHEN / OUR REF.

P 28 241-00975/co

DATUM / DATE

02.10.1996

Prof. Dr. Hermann Bujard
Remlerstr. 9
69120 Heidelberg

Rekombinantes Herstellungsverfahren für ein vollständiges Malaria-Antigen gp190/MSP1

Rekombinantes Herstellungsv erfahren für ein vollständig s Malaria-Antigen gp190/MSP1

Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das vollständige Malaria-Antigen gp190/MSP1 sowie einzelner natürlich vorkommender Domänen und Teile derselben durch Expression (einer) synthetischer DNA-Sequenzen. Die Erfindung betrifft außerdem die durch das Verfahren hergestellten DNA-Sequenzen und die für die Expression der DNA-Sequenzen verwendeten Wirtsorganismen. Ferner betrifft die Erfindung die Verwendung des vollständigen Malaria-Antigens sowie Teile derselben als Impfstoff zur Immunisierung gegen Malaria.

Schließlich betrifft die vorliegende Erfindung ein Stabilisierungsverfahren für AT-reiche Gene, sowie stabilisierte Gene, die sich durch einen geringeren AT-Gehalt auszeichnen.

Malaria ist weltweit eine der bedeutendsten Infektionskrankheiten. Nach Angaben der WHO waren 1990 in 99 Ländern 40% der Weltbevölkerung einem Malaria-Risiko ausgesetzt. Ihre Verbreitung nimmt derzeit wieder massiv zu. Dies ist vor allem auf eine intensive Resistenzbildung der Malariaerreger zurückzuführen, die dadurch gefördert wird, daß die zur Therapie eingesetzten Medikamente ebenfalls als Prophylaxe empfohlen und eingenommen werden. Neben der Suche nach neuen, wirksamen Chemothrapeutika werden heute Hoffnungen auch in die Entwicklung von Impfstoffen gesetzt, da Menschen in Malaria-epidemischen Regionen der Welt verschiedene Arten von Immunität zu entwickeln vermögen. Neben einer natürlichen Resistenz gegen Malaria, die sich bei heterozygoten Trägern des Sichelzellgens und bei Personen mit Thalassämie und Glucose-6-Phosphat-Dehydrogenase-Mangel ausbildet, können im Laufe einer Malariainfektion im Menschen Immunitätsmechanismen einsetzen, die sich in einer erhöhten Widerstandsfähigkeit gegenüber den Plasmodien äußern. Dementsprechend ist der Krankheitsverlauf in stark durchseuchten Bevölkerungspopulationen weitaus weniger bedrohlich als bei

Personen, die der Infektion weniger häufig oder erstmalig ausgesetzt sind.

Das Hauptproblem bei der Impfstoffentwicklung ist die Identifikation eines Antigens, das schützende Immunität bewirken kann, da kein leicht zugängliches und gut definiertes Tiermodell für die vier den Menschen befallenden Parasiten vorhanden ist. Die Malaria-Erreger gehören der Gruppe der Plasmodien an, wobei die Infektion mit einem der vier Erreger *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae* oder *Plasmodium falciparum* durch den Stich von Anopheles-Mücken erfolgt. Von diesem Erreger ist *Plasmodium falciparum* der gefährlichste und der am weitesten verbreitete.

Das Hauptoberflächenprotein von Merozoiten, der invasiven Form der Blutstadien des Malaria-Erregers *Plasmodium falciparum* und anderer Malaria-Erreger wie *P. vivax*, ist ein 190 - 220 kD Glykoprotein. Spät in der Entwicklung des Parasiten wird dieser Vorläufer in kleinere Proteine prozessiert, die jedoch als einheitlicher Komplex aus Merozoiten isoliert werden können. Der Komplex ist mittels eines Glycosylphosphatidyl-Inositol-Ankers mit der Merozoitenmembran verknüpft. Die Sequenzen der gp 190-Proteine verschiedener *P. falciparum*-Stämme fallen in zwei Gruppen, zwischen denen intragene Rekombination häufig ist. Insgesamt besteht das Protein aus mehreren hochkonservierten Regionen, aus einem dimorphen Bereich, welche jeweils einem von zwei Allelen angehören und aus zwei relativ kleinen oligomorphen Blöcken im N-terminalen Bereich (Tanabe, K., Mackay, M., Goman, M. und Scaife, J.G. (1987), Allelic dimorphism in a surface antigen gene of the malaria parasite *Plasmodium falciparum*. *J. Mol. Biol.* 195, 273-287; Miller, L.H., Roberts, T., Shahabuddin, M. und McCutchan, T.F. (1993), Analysis of sequence diversity in the *Plasmodium falciparum* merozoite surface protein-1 (MSP-1). *Mol. Biochem. Parasitol.* 59, 1-14).

Das gp190/MSP1 galt bereits früh als ein möglicher Kandidat für einen Impfstoff. So wurde im Nagermodell nach Immunisierung mit

dem analogen Protein aktiver Schutz gegen Infektion mit Nager-Parasiten erhalten. Passiver Schutz ließ sich mit gegen dieses Protein gerichteten Antikörpern erreichen (siehe auch Holder, A. A. und Freeman, R.R. (1981), Immunization against blood-stage rodent malaria using purified parasite antigens, *Nature* 294, 361-364; Majarian, W.R., Daly, T. M., Weidanz, W.P. und Long, C.A. (1984), Passive immunization against murine malaria with an IgG3 monoclonal antibody, *J. Immunol.* 132, 3131-3137). Die Daten, die diese Annahme belegen sollen, sind im einzelnen jedoch statistisch nicht signifikant.

Darüber hinaus sind mehrere monoklonale Antikörper, welche *in vitro* die Invasion von Erythrozyten durch *P. falciparum* inhibieren, sind gegen gp190/MSP1 gerichtet (Pirson, P.J. und Perkins, M.E. (1985), Characterization with monoclonal antibodies of a surface antigen of *Plasmodium falciparum* merozoites. *J. Immunol.* 134, 1946-1951; Blackman, M.J., Heidrich, H.-G., Donachie, S., McBride, J.S. und Holder A.A. (1990), A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies, *J. Exp. Med.* 172, 379-382).

Schließlich wurde eine Reihe von Impfstudien mit gp190/MSP1-Material aus *P. falciparum* an Primaten, insbesondere an Aotus und Saimiri-Affen durchgeführt (siehe auch Perrin, L.H., Merkli, B., Loche, M., Chizzolini, C., Smart, J. und Richle, R. (1984), Antimalarial immunity in Saimiri monkeys. Immunization with surface components of asexual blood stages, *J. Exp. Med.* 160, 441-451; Hall, R., Hyde, J.E., Goman, M., Simmons, D.L., Hope, I.A., Mackay, M. und Scaife, J.G. (1984), Major surface antigen gene of a human malaria parasite cloned and expressed in bacteria, *Nature* 311, 379-382; Siddiqui, W.A., Tam, L.Q., Kramer, K.J., Hui, G.S.N., Case, S.E., Yamaga, K.M., Chang, S.P., Chan, E.B.T. und Kan, S.-C. (1987), Merozoite surface coat precursor protein completely protects Aotus monkeys against *Plasmodium falciparum* malaria, *Proc. Natl. Acad. Sci. USA* 84, 3014-3018; Ettlinger, H.M., Caspers, P., Materile, H.,

Schoenfeld, H.-J., Stueber, D. und Takacs, B. (1991), Ability of recombinant or native proteins to protect monkeys against heterologous challenge with *Plasmodium falciparum*, Inf. Imm. 59, 3498-3503; Holder, A.A., Freeman, R.R. und Nicholls, S.C. (1988), Immunization against *Plasmodium falciparum* with recombinant polypeptides produced in *Escherichia coli*, Parasite Immunol. 10, 607-617; Herrera, S., Herrera, M.A., Perlaza, B.L., Burki, Y., Caspers, P., Döbeli, H., Rotmann, D. und Certa, U. (1990), Immunization of Aotus monkeys with *Plasmodium falciparum* blood-stage recombinant proteins, Proc. Natl. Acad. Sci. USA 87, 4017-4021; Herrera, M.A., Rosero, F., Herrera, S., Caspers, P., Rotmann, D., Sinigaglia, F. und Certa, U. (1992), Protection against malaria in Aotus monkeys immunized with a recombinant blood-stage antigen fused to a universal T-cell epitope; correlation of serum gamma interferon levels with protection, Inf. Imm. 60, 154-158; Patarroyo, M.E., Romero P., Torres, M.L., Clavijo, P., Moreno, A., Martinez, A., Rodriguez, R., Guzmann, F. und Cabezas, E. (1987), Induction of protective immunity against experimental infection with malaria using synthetic peptides, Nature 328, 629-632). In diesen Impfstudien lassen sich hierbei zwei Ansätze unterscheiden

- Verwendung von aus Parasiten isoliertem Material

und

- Einsatz von in heterologen Expressionssystemen gewonnenem Material.

Letzteres bestand in der Regel aus relativ kleinen Teilbereichen des Gesamtproteins. Obwohl die ersten Resultate der in Vorversuchen an Affen durchgeföhrten Impfungen andeuten, daß gp190/MSP1 einen Schutz vermitteln könnte, haben alle an Primaten durchgeföhrten Experimente zwei Probleme, welche einen solchen Schluß in Frage stellen:

(a) sie wurden an zu kleinen Tiergruppen durchgeföhr

(b) sie wurden nicht wiederholt.

Die Resultate und die daraus gezogenen Schlüsse sind daher statistisch nicht abgesichert. Neben dem schwierigen Zugang zu geeigneten Affen liegt das zugrunde liegende Hauptproblem darin, daß es bislang nicht möglich war, gutes Impfmaterial in ausreichender Menge herzustellen.

Andererseits konnten nach der Sequenzierung des gp190-Gens aus dem K1- und dem MAD20-Stamm von *Plasmodium falciparum* konnten überlappende Fragmente in *E. coli* exprimiert werden. Mit diesem Material zeigten epidemiologische Studien in Westafrika, daß in der Gruppe der Adoleszenten eine Korrelation bestand zwischen Antikörpertiter gegen gp190/MSP1-Fragmente einerseits und einem Schutz vor Parasiteninfektion andererseits. Darüber hinaus schien der Titer auch mit der Fähigkeit zu korrelieren, die Parasitämie auch auf niedrigem Niveau zu kontrollieren (Tolle et al. (1993): A prospective study of the association between the human humoral immune response to *Plasmodium falciparum* blood stage antigen gp190 and control of malarial infections. Infect Immun. 61, 40-47). Diese Resultate werden ergänzt durch neue Untersuchungen an Aotusaffen im Rahmen der vorliegenden Erfindung. Hier wurde ein hoher Schutz gegen Infektion mit dem Parasiten dadurch erreicht, daß Proteinpräparationen aus *Plasmodium falciparum*, die überwiegend aus nichtprozessiertem gp190/MSP1 bestanden, als Impfstoff benutzt worden waren. Die Affen mit dem höchsten Antikörpertiter gegen gp190/MSP1 waren am besten geschützt. Diese Resultate machten letztendlich das gp190 zu einem vielversprechenden Kandidaten für eine Impfstoff gegen Malaria tropica.

Von einigen Arbeitsgruppen wurde der C-terminalen Domäne des gp190 (p19 bzw. p42) eine besondere Rolle bei der gp190 vermittelten Immunität zugewiesen (siehe auch Chang, S.P., Case, S.E., Gosnell, W.L., Hashimoto, A., Kramer, K.J., Tam, L. Q., Hashiro, C.Q., Nikaido, C.M., Gibson, H.L., Lee-Ng, C.T., Barr,

P.J., Yokota, B.T. und Hui, G.S.N. (1996), A recombinant baculovirus 42-kilodalton C-terminal fragment of *Plasmodium falciparum* merozoite surface protein 1 protects Aotus monkeys against malaria, *Inf. Imm.* 64, 253-261; Burghaus, P.A., Wellde, B.T., Hall, T., Richards, R.L., Egan, A.F., Riley, E.M., Ripley-Ballou, W. und Holder A.A. (1996), Immunization of Aotus nancymai with recombinant C-terminus of *Plasmodium falciparum* merozoite surface protein-1 in liposomes and alum adjuvant does not induce protection against a challenge infection, *Inf. Imm.*, in press).

Bislang ist es jedoch nicht möglich, auf rationaler Basis andere Teile des gp190 als nicht relevant für eine schützende Immunantwort auszuschließen. Es ist daher nach wie vor notwendig, das gesamte Gen bzw. das intakte gp190 für Impfversuche zu verwenden. Trotz mehrfacher Versuche verschiedener Arbeitskreise ist es jedoch noch nicht gelungen, das gesamte Gen des gp190/MSP1 zu klonieren und zu exprimieren.

Bislang war es auch noch nicht möglich, a priori einen Teil aus der gp190-Sequenz für die schützende Immunantwort als nicht relevant auszuschließen, so daß es nach wie vor notwendig ist, das gesamte Gen bzw. das gesamte Genprodukt für Impfversuche zu verwenden. Es ist jedoch trotz vieler Versuche mehrerer Arbeitskreise noch nicht gelungen, das gesamte Gen des gp190/MSP1 zu klonieren.

Es war daher eine Aufgabe der vorliegenden Erfindung, Impfmateri al in Form von vollständigem gp190/MSP1 in ausreichender Menge zur Verfügung zu stellen. Es war eine weitere Aufgabe der vorliegenden Erfindung, ein Verfahren anzugeben, mit dem dieses Impfmateri al gewonnen werden kann.

Es war außerdem eine weitere Aufgabe gemäß der vorliegenden Erfindung, eine vollständige DNA-Sequenz von gp190/MSP1 anzugeben, die in einem Wirtsorganismus exprimierbar ist.

Weiterhin war es eine Aufgabe der vorliegenden Erfindung, Wirtsorganismen anzugeben, die das vollständige Gen gp190/MSP1 enthalten.

Schließlich war es auch eine Aufgabe der vorliegenden Erfindung, ein Stabilisierungsverfahren für AT-reiche Gene anzugeben, sowie ein für die Expression geeignetes, stabilisiertes Gen, das sich durch eine Erniedrigung des AT-Gehalts auszeichnet.

Diese Aufgaben werden durch die in den Ansprüchen angegebenen Gegenstände gelöst.

Im folgenden werden einige Begriffe näher erläutert, um klarzustellen, wie sie in diesem Zusammenhang verstanden werden sollen.

"Rekombinantes Herstellungsverfahren" bedeutet, daß ein Protein von einer DNA-Sequenz durch einen geeigneten Wirtsorganismus exprimiert wird, wobei die DNA-Sequenz aus einer Klonierung und Fusion einzelner DNA-Abschnitte entstanden ist.

"Vollständiges gp190/MSP1-Protein" meint hier das gesamte, aus o.g. Plasmodien, insbesondere *Plasmodium falciparum*, isolierbare gp190/MSP1-Oberflächenprotein, das das Hauptoberflächenprotein der Merozoiten des o.g. Erregers darstellt sowie die Proteine mit analoger Funktion aus den anderen Plasmodiumarten, wie *P. vivax*. Der Begriff betrifft somit jeweils das Hauptoberflächenprotein der Merozoiten der 4 o.g. für den Menschen gefährlichen Malariaerreger. "Vollständiges gp190/MSP1-Gen" ist das für dieses Protein kodierende Gen. "Vollständig" bedeutet in diesem Zusammenhang, daß die gesamte Aminosäuresequenz des nativen Proteins vorhanden ist, bzw. daß die Gensequenz für die gesamte Aminosäuresequenz des nativen Proteins kodiert. Mit umfaßt sind jedoch auch mutierte und/oder verkürzte Formen von gp190/MSP1, sofern sie das gleiche Immunisierungspotential (Impfschutz) wie das vollständige gp190/MSP1 aufweisen.

Schließlich umfaßt der Begriff auch Varianten von gp190/MSP1 die sich dadurch auszeichnen, daß sie Proteinabschnitte verschiedener Allele in einem Proteinmolekül enthalten.

"FCB-1" ist ein Stamm von *P. falciparum*, wie beschrieben bei Heidrich, H.-G., Miettinen-Baumann, A., Eckerskorn, C. und Lottspeich, F. (1989) The N-terminal amino acid sequences of the *Plasmodium falciparum* (FCB1) merozoite surface antigens of 42 and 36 kilodalton, both derived from the 185-195-kilodalton precursor. *Mol. Biochem. Parasitol.* 34, 147-154.

"Ankersignal" meint hier eine Proteinstruktur, für die eine DNA-Sequenz am 3'- oder 5'-Ende des erfindungsgemäßen Gens kodiert. Ankersignale sind Strukturen, die einem Polypeptid die Verankerung an anderen Strukturen, wie z.B. Membranen ermöglichen.

"Signalpeptid" bedeutet hier eine Proteinstruktur, für die eine DNA-Sequenz am N-terminalen Ende des erfindungsgemäßen Gens kodiert. Signalpeptide sind Strukturen, die dem Polypeptid u.a. ein Einschleusen in Membranen ermöglichen.

"AT-Gehalt" meint im Zusammenhang der vorliegenden Erfindung die prozentuale Menge von Adenin/Thymin-Basenpaaren im Verhältnis zu Guanin/Cytosin-Basenpaaren.

"Klonierung" soll hier alle im Stand der Technik bekannten Klonierungsmethoden umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen beschrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

"Expression in einem geeigneten Expressionssystem" soll hier alle im Stand der Technik bekannten Expressionsmethoden in bekannten Expressionssystemen umfassen, die hier zum Einsatz kommen könnten, die jedoch nicht alle im einzelnen beschrieben werden, weil sie zum selbstverständlichen Handwerkszeug des Fachmanns gehören.

Es ist eine erste Aufgabe gemäß der vorliegenden Erfindung, ein Verfahren anzugeben, durch das das gp190/MSP1-Protein und das Gen hierfür in ausreichender Menge ohne übermäßige Kosten produziert werden kann.

Diese Aufgabe wird durch das in Anspruch 1 beschriebene rekombinante Herstellungsverfahren gelöst, durch das ein vollständiges gp190/MSP-1-Gen und das davon kodierte Protein in ausreichender Menge erhältlich ist.

Durch dieses Verfahren wurde es erstmals möglich, das Protein in seiner Gesamtheit außerhalb des Parasiten zu synthetisieren. Das so synthetisierte Protein ist, wie die Analyse mit konformationelle Epitope erkennenden monoklonalen Antikörpern zeigt, zumindest über weite Bereiche in natürlich gefalteter Form herstellbar. Durch das rekombinante Herstellungsverfahren konnten jeweils mehrere Milligramm intaktes gp190/MSP1 aus den Wirtsorganismen gewonnen werden, eine Menge, die aus technischen und aus Kostengründen aus Parasiten praktisch nicht gewonnen werden kann. Die jetzt mögliche Produktion des Proteins in beliebigen Mengen eröffnet neue Perspektiven für seinen Einsatz als experimentellen Impfstoff gegen Malaria. Darüber hinaus ist der Weg frei für die Entwicklung von Lebendimpfstoffen sowie für Vakzine auf Nukleinsäure-Basis.

Vorzugsweise liegt der Synthese der für das Protein gp190/MSP1 kodierenden Gen-Sequenz die Sequenz des *P. falciparum* FCB-1 Stammes zugrunde. *P. falciparum* ist der Erreger der Malaria tropica und damit der gefährlichste unter den Malaria-Arten. Das zugrunde liegende Gen ist ein Vertreter des "K1-Allels", wobei K1 für einen bestimmten *P. falciparum*-Stamm steht. Seine kodierende Sequenz erstreckt sich über 4917 Basenpaare und schließt eine Signalsequenz am N-terminalen Ende sowie eine Anker-Sequenz am C-terminalen Ende ein.

Weiterhin ist das rekombinante Herstellungsverfahren gemäß der Erfindung vorzugsweise dadurch gekennzeichnet, daß der AT-

Gehalt der dem Protein zugrunde liegenden DNA-Sequenz gegenüber der Wildtyp-Sequenz erniedrigt ist, von 74% im ursprünglichen Gen auf vorzugsweise ca. 55%, indem bspw. unter Erhalt der Aminosäuresequenz des FCB-1-Proteins eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten hergestellt wird. Andere Codonhäufigkeiten, welche den AT-Gehalt erniedrigen, sind ebenfalls denkbar.

Vorzugsweise kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und GPI-Ankersignalpeptid, im weiteren als gp190^s bezeichnet.

In einer weiteren bevorzugten Ausführungform kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des GPI-Ankersignals. Diese Ausführungsform wird im weiteren als gp190^{s1} bezeichnet.

In noch einer weiteren bevorzugten Ausführung kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des GPI-Ankersignals und des Signalpeptids. Diese Ausführungsform wird im weiteren als gp190^{s2} bezeichnet.

In einer weiteren bevorzugten Ausführungsform kodiert das dem durch das rekombinante Herstellungsverfahren erzeugte Protein zugrunde liegende Gen für die volle Aminosäuresequenz und eine Transmembranankersequenz.

In einer besonders bevorzugten Ausführungsform umfaßt das rekombinante Herstellungsverfahren folgende Schritte.

Zunächst den Entwurf der zu synthetisierenden DNA-Sequenz auf der Basis des Gens aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den z.B. im menschlichen Genom üblichen Codon-

Häufigkeiten unter Beibehalten der Aminosäuresequenz des FCB-1-Proteins hergestellt wird.

Hierdurch sollte der AT-Gehalt des Gens reduziert werden, vorzugsweise auf 55%.

Im weiteren Verfahren wird die entworfene Sequenz bspw. in 5 überlappende Regionen eingeteilt, welche jeweils Domänen der natürlichen Prozessierungsprodukte des gp190/MSP1-Proteins aus FCB-1 entsprechen: p83, p31, p36, p30 und p19.

Es werden Desoxyoligonukleotide synthetisiert, die jeweils mindestens die gesamte Länge einer Region abdecken.

Besonders bevorzugt werden die Desoxyoligonukleotide so synthetisiert, daß ihre Sequenz abwechselnd dem "oberen" (5' - 3') bzw. dem "unteren" (3' - 5') DNA-Strang entspricht. Die Länge dieser Oligonukleotide ist vorzugsweise durchschnittlich 120 Nukleotide und sie überlappt die benachbarten Sequenzen jeweils um ca. 20 Basen.

In einer möglichen Ausführungsform werden DNA-Sequenzen von etwa doppelter Länge wie die jeweiligen Ausgangsprodukte durch asymmetrische PCR hergestellt und zwar so, daß die überschüssigen DNA-Sequenzen, die benachbart sind, jeweils den gegenüberliegenden Strang repräsentieren. Dies führt in einem zweiten

PCR-Amplifikationszyklus zu einem Zweitprodukt, das der Länge von vier ursprünglich eingesetzten Oligonukleotiden (abzüglich der überlappenden Region) entspricht. Die Überführung dieser Produkte in ein überwiegend aus Einzelstrang-DNA bestehendes Präparat durch asymmetrische PCR mit den endständigen Oligonukleotiden erlaubt in einem weiteren Amplifikationsschritt die Herstellung eines 800 bp langen doppelsträngigen DNA-Fragments in nur 25 PCR-Zyklen.

Auf diese Weise werden direkt die kodierenden Regionen für p19, p30, p36 und p31 synthetisiert und in *E. coli* molekular kloniert. Klone mit fehlerfreier Sequenz wurden entweder direkt oder durch Zusammensetzen fehlerfreier Teilsequenzen erhalten. Die Region, welche das p83 kodiert, wurde durch Fusion aus zwei etwa 1200 bp umfassenden Sequenzen erhalten.

Im weiteren Verlauf des Verfahrens wurden die einzelnen Sequenzen kloniert. Als Expressionsvektoren bieten sich vorzugsweise die Plasmide pDS56, RBSII ("Hochuli, E., Bannwarth, W. Döbeli, H. Gentz, R., and Stüber, D. (1988) Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Biotechn. 6, 1321-1325"), pBi-5 ("Baron, U., Freundlieb, S., Gossen, M. and Bujard, H. (1995) Co-regulation of two gene activities by tetracycline via a bidirectional promoter. Nucl. Acids Res. 23, 3605-3606") und ppTMCS an. Es sind jedoch auch andere Expressionsvektoren denkbar.

Bevorzugte Wirtsorganismen für die Expression sind *E. coli*, besonders bevorzugt der Stamm DH5alphaZ1 (R. Rutz, Dissertation 1996, Universität Heidelberg), HeLa-Zellen, CHO-Zellen, *Toxoplasma gondii* (Pfefferkorn, E.R. and Pfefferkorn, C.C. 1976, *Toxoplasma gondii: Isolation and preliminary characterization of temperature - sensitive mutants. Exp. Parasitol.* 39, 365-376) oder *Leishmania*. Weitere Wirtssysteme könnten z.B. Hefen, Baculoviren oder Adenoviren sein, wobei der Gegenstand der Erfindung nicht auf die genannten Wirtssysteme beschränkt sein sollte.

Es war eine weitere Aufgabe der vorliegenden Erfindung, eine vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von *P. falciparum* anzugeben.

Diese Aufgabe wird durch die in Anspruch 17 genannte Erfindung gelöst, wobei die Sequenz durch das oben beschriebene rekombinante Herstellungsverfahren erhältlich ist.

In einer bevorzugten Ausführungsform der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz.

In einer anderen bevorzugten Ausführungsform der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz mit Ausnahme des Ankersignals.

In einer weiteren bevorzugten Ausführungsform gemäß der vorliegenden Erfindung kodiert die zur Expression geeignete DNA-Sequenz für die vollständige Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptids. Diese Ausführungsform von gp190/MSP1 kann dadurch gekennzeichnet sein, daß sie am N-Terminus 11 zusätzliche Aminosäuren, davon 6 Histidine, enthält.

Besonders bevorzugt enthält die zur Expression geeignete DNA-Sequenz keine erkennbaren "splice-donor" und "splice-acceptor"-Signale, und sie ist vorzugsweise dadurch gekennzeichnet, daß sie keine größeren GC-reichen Sequenzen enthält, die stabile Haarnadelstrukturen auf RNA-Ebene bewirken könnten.

Vorzugsweise sollten Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs und mehr Basenpaaren erkennen, vermieden werden.

In einer bevorzugten Ausführung werden spezifische, d.h. nur einmal im Gen vorkommende Schnittstellen für Restriktionsendonukleasen in Regionen eingeführt, welche die nach der Prozesierung des Proteins entstehenden Domänen trennen.

Besonders bevorzugt sollten an beiden Enden des Gens Sequenzen für Restriktionsendonukleasen vorhanden sein, die im Gen nicht vorkommen.

Weiterhin werden durch die Erfindung Wirtsorganismen zur Verfügung gestellt, die die vollständige Sequenz des gp190/MSP-1 Oberflächenproteins enthalten.

Solche Wirtsorganismen sind vorzugsweise *E. coli*, besonders bevorzugt der Stamm DH5alphaZ1, HeLa-Zellen, CHO-Zellen, *Toxoplasma gondii* oder *Leishmania*. Die HeLa-Zellen und CHO-Zellen sollten vorzugsweise konstitutiv tTA synthetisieren.

Schließlich stellt die vorliegende Erfindung eine Möglichkeit zur Verfügung, ein nach dem rekombinanten Herstellungsverfahren erzeugtes gp190/MSP1-Oberflächenprotein oder Teile desselben zur aktiven Immunisierung gegen Malaria zu verwenden.

Das hier vorgestellte Syntheseschema erlaubt auch das zweite Allel des gp190/MSP1-Gens herzustellen. Damit wird dem Dimorphismus des Proteins Rechnung getragen. Die Hauptvariabilität des Proteins beruht jedoch auf den Sequenzen von zwei relativ kurzen Blöcken (Block II und IV, Ref. 1), die oligomorph sind. Die vorliegenden Sequenzdaten ermöglichen es, daß mit 6-8 Sequenzkombinationen dieser Blöcke über 95% aller bekannten gp190/MSP1-Sequenzen abgedeckt werden können. Die Synthese dieser Sequenzvarianten läßt sich problemlos anhand der hier vorgestellten Strategien verwirklichen, so daß Varianten sowohl in das K1 als auch in das MAD20-Allel eingebaut werden können. Impfstoffe aus den hierdurch entstehenden Sequenzfamilien können ggf. gegen ein breites Spektrum von Parasiten mit gp190/MSP1-Varianten Schutz verleihen.

Die Herstellung von verschiedenen Impfstofftypen ist möglich:

- Auf der Ebene von Proteinpräparaten, wobei jeweils Mischungen der zwei Familien (K1-Typ, MAD20-Typ mit verschiedenen Varianten der Blöcke II und IV) zur Anwendung kommen können. Verschiedene Träger bzw. Adjuvans-Materialien können zum Einsatz kommen: Aluminiumoxid, Liposomen, Iscoms QSz1 etc.

- Auf der Ebene der Lebendimpfstoffe: (a) virale Träger, insbesondere Vakzinia und Adenoviren; (b) Parasiten als Träger, insbesonder avirulente Formen von Leishmania und Toxoplasma; (c) bakterielle Träger, z.B. Salmonella
- Auf der Ebene der Nukleinsäuren, wobei bspw. für die Gentherapie geeignete Vektoren als Vehikel zum Einbringen der Gene in den Wirt verwendet werden; weiterhin ist das Einbringen von Ribonukleinsäuren, welche das gewünschte Protein kodieren, denkbar.

Eine weitere Möglichkeit der Impfung besteht in der Verwendung eines gemäß dem erfindungsgemäßen rekombinanten Herstellungsverfahren erzeugten gp190/MSP1-Proteins zur Herstellung von monoklonalen Antikörpern, die dann ihrerseits zur passiven Immunisierung gegen Malaria verwendet werden.

Ebenso wird eine Verwendung der gemäß dem rekombinanten Herstellungsverfahren in einem Zwischenschritt entstandenen, dem Protein zugrunde liegenden DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinsäurebasis ermöglicht.

Schließlich betrifft die Erfindung auch ein Verfahren zur Stabilisierung von Gen-Sequenzen, insbesondere von Sequenzen, die keine ausreichende Stabilität in Expressionssystemen zeigen.

Diese Stabilisierung wird erfindungsgemäß dadurch erreicht, daß der AT-Gehalt der Sequenz verringert wird.

Weiterhin wird durch die Erfindung ein stabilisiertes Gen zur Verfügung gestellt, das sich dadurch auszeichnet, daß es einen geringeren AT-Gehalt aufweist. Ein Beispiel für ein solches, stabilisiertes Gen ist das Gen für das gp190/MSP1-Oberflächenprotein gemäß der vorliegenden Erfindung.

Im folgenden soll die Erfindung anhand der Abbildungen und Tabellen sowie einiger Beispiele in einzelnen Ausführungsformen beschrieben werden.

Dabei zeigt:

Abb. 1: Schematische Darstellung des gp190/MSP1 Vorläuferproteins aus *P. falciparum* (FCB-1).

Abb. 2: Zwei mit nativem gp190/MSP1 aus *P. falciparum* (FCB-1) an Aotus-Affen durchgeführte Impfversuche.

Abb. 2a: mit 3 x 60 Mikrogramm gp190/MSP1

Abb. 2b: mit 3 x 40 Mikrogramm gp190/MSP1

Abb. 3A: Strategie der Synthese des gp190/MSP1-Gens

Abb. 3B: Prinzip der PCR-gestützten Total-Synthese

Abb. 3C: Totalsequenz des gp190^S

Abb. 3D: N- und C-Terminus der gp190^{S1}-Variante

Abb. 4A: Expressionsvektor pDS56 mit gp190^{S2}-Sequenz

Abb. 4B: Gelelektrophorese von gp190^{S2}.

Abb. 5A: Expressionsvektor pBi-5 mit gp190^{S1}-Sequenz

Abb. 5B: Immunfluoreszenz von HeLa-Zellen

Abb. 5C: Elektrophoretische Charakterisierung von aus HeLa-Zellen aufgereinigtem gp190^{S1}

Abb. 6A: Expressionsvektor ppT 190 mit gp190-Sequenz

Abb. 6B: Immunfluoreszenz der Expression von gp190^S in T. gondii

Abb. 6C: Polyacrylamid-Gelelektrophorese von gp190 aus T. gondii

Bei dem in Abb. 1 schematisch dargestellten gp190/MSP1-Vorläuferprotein aus *P. falciparum* (FCB-1) repräsentieren die dunklen Blöcke Regionen, die in allen Stämmen hochkonserviert vorliegen. Die schraffierten Blöcke zeigen die dimorphen Bereiche, welche im Falle des FCB-1 Isolates dem K1-Allel entstammen. O1 und O2 zeigen die oligomorphen Bereiche. S bezeichnet die Signalpeptidsequenz, welche 19 Aminosäuren umfaßt, GA, die C-terminale Region, welche das Signal für die GPI-Verankerung des Proteins in der Membran enthält. Die Pfeile deuten die Prozessierungsstellen an, durch die Proteine p53, p31, p36, p30 und p19 entstehen. Das gp190-Gen kodiert insgesamt 1639 Aminosäuren.

Die weiteren Abbildungen werden im Zusammenhang mit den folgenden Beispielen ausführlicher erläutert.

BEISPIELE

Beispiel 1: Totalsynthese einer das gp190/MSP1 kodierenden DNA-Sequenz (siehe hierzu Abb. 3)

A. Strategie der Synthese des gp190/MSP1-Gens (gp190^S) (siehe Abb. 3A).

Die Sequenz wurde aufgeteilt in Fragmente, die den Hauptprozessierungsprodukten entsprachen: p83, p31, p36, p30 und p19. In den Übergangsregionen wurden Spaltstellen für Restriktionsendonukleasen (Pfeile in Abb. 3) so eingeplant, daß die Aminosäuresequenz nicht verändert wurde. Alle hier aufgeführten Schnittstellen kommen in der Sequenz nur einmal vor.

Die Fragmente wurden überlappend synthetisiert, so daß die Schnittstellen an den jeweiligen Enden die Verknüpfung zu den benachbarten Fragmenten durch Ligierung möglich machten. Alle Einzelfragmente enthielten zusätzlich an ihrem 5'-Ende eine BamHI-Schnittstelle zur Insertion in Expressionsvektoren. Über MluI und ClaI konnte die Gesamtsequenz kloniert werden. Das hier gezeigte Schema führt zunächst zu einer Sequenz, welche den GPI-Anker nicht ausbilden kann, da C-terminal 18 Aminosäuren fehlen. Die Synthese eines entsprechenden Oligonukleotids, sowie eines über die SphI-Schnittstelle sich erstreckenden "Primers" führt nach PCR zu dem Fragment GA, das über SphI und ClaI eingesetzt werden konnte, die resultierende Totalsequenz war gp190^S. Zur Entfernung der das Signalpeptid kodierenden Sequenz wurden "PCR-Primer" hergestellt, über die das Fragment ΔS synthetisiert wurde. Es erlaubt, über eine BamHI und eine HindIII-Schnittstelle den N-Terminus so zu verändern, daß das Protein mit Aminosäure Nr. 20 begann. Die Kernsequenz, welche das gp190/MSP1 ohne Signalsequenz und ohne GPI-Anker-Signal kodiert, wurde mit gp 190^{S2} bezeichnet. Deletion des GPI-Ankersignals allein führte zu gp190^{S1}.

B. Prinzip der PCR-gestützten Totalsynthese (siehe Abb. 3B)

Oligodesoxynukleotide von etwa 120 Nukleotiden Länge wurden abwechselnd vom kodierenden bzw. nichtkodierenden Strang so synthetisiert, daß sie jeweils etwa 20 Basen mit dem benachbarten Fragment überlappten. Das Schema zeigt beispielhaft die Synthese eines ca. 800 bp langen Fragments aus Oligonukleotiden. In der ersten Stufe wurden in 4 Reaktionsgefäßern jeweils 2 Oligonukleotide "asymmetrisch" amplifiziert. Es entstanden 4 etwa 220 bp lange DNA-Populationen, die vorwiegend aus Einzelsträngen bestanden (A, B, C, D). Vereinigung von A und B sowie von C und D und Amplifikation über 5 Zyklen führte zu 2 etwa 400 bp langen doppelsträngigen Produkten. Asymmetrische Amplifikation dieser DNA-Fragmente (Stufe III) ergab Einzelstrangpopulationen, welche nach Vereinigung und Amplifikation (Stufe IV) nach

10 Zyklen das Endprodukt G von ca. 800 bp Länge ergaben. Diese Synthese war ohne Isolierung der Zwischenprodukte und ohne Puffer oder Enzymerneuerung durchführbar, und war nach 3 Stunden beendet. Das Endprodukt wurde elektrophoretisch gereinigt, mit den geeigneten Restriktionsendonukleasen nachgeschnitten und im pBluescript (Stratagene), in dessen Polylinker eine MluI und eine ClaI-Schnittstelle eingesetzt worden waren, in *E. coli* kloniert.

C. Totalsequenz des gp190^S (siehe Abb. 3C)

Nach Fusion aller Teilsequenzen (Abb. 3A) in pBluescript wurde die Sequenz des Gens mit der Dideoxymethode verifiziert. Das Leseraster des gp190^S hatte eine Länge von 4917 bp (+2 Stopcodons) und kodierte eine Aminosäuresequenz, welche der des gp190/MSP1 aus FCB-1 entspricht (1639 Aminosäuren).

D. N- und C-Terminus der gp190^{S1}-Variante (siehe Abb. 3D)

Die N-terminale Sequenz, beginnend mit der BamHI-Schnittstelle, zeigte den Übergang bei Aminosäure 20, von der angenommen wird, daß sie nach Abspaltung des Signalpeptids den N-Terminus definiert. Am C-Terminus war die kodierte Sequenz bei Aminosäure 1621 zu Ende. Den Stopcodons folgte die ClaI-Schnittstelle.

Beispiel 2: Expression des gp190^{S2} in *E. coli*

A. Expressionsvektor (siehe Abb. 4A)

Die gp190^{S2}-Sequenz wurde über die BamHI- und ClaI-Schnittstellen in pDS56RBSII eingesetzt. Dadurch wurden 6 Histidine sowie einige aus dem Vektor stammenden Aminosäuren an den N-Terminus fusioniert; dies ergibt folgende N-terminale Sequenz des Leserasters; Met Arg Gly Ser (His)₆ Gly Ser. Durch den Promotor P_{N251aco-1} stand die Transkription unter lac R/O/IPTG-Kontrolle.

B. Expression und Aufreinigung von gp190^{S2} (siehe Abb. 4D)

Eine Überführung des Vektors pDS56RBSIIgp190^{S2} in E. coli DH5 α Z1 und Induktion der Synthese durch IPTG ergab nach elektrophoretischer Auftrennung des Protein-Totalextraktes der Kultur eine deutlich sichtbare Bande in der erwarteten Größe (Pfeil). Die Aufreinigung des Materials durch IMAC und Affinitätschromatographie (Antikörpersäule mit mAK5.2) führte zu einem homogenen Produkt von etwa 190 kD. In der Abb. bedeuten M= Molekulargewichtsstandards; 1= E. coli Proteine vor, 2= nach Induktion mit IPTG für 2 Stunden. 3,4,5 = Fraktionen aus der Elution der mAK-Säule.

Beispiel 3: Tetrazyklin-kontrollierte Expression von gp190^{S1} in HeLa- und CHO-Zellen und Isolation des Produktes (siehe auch Abb. 5) bzw. 6c)

A. Die gp190-Sequenz wurde in den Expressionsvektor pBi-5 über die BamHI/Clal-Schnittstellen eingesetzt. Damit stand die Transkription des Gens unter Kontrolle eines bidirektionalen "tTA-responsive"-Promotors und konnte über Tc reguliert werden. Der bidirektionale Promotor initiierte gleichzeitig die Transkription des Indikatorgens Luciferase. Damit ließ sich die Regulation der Expression leicht verfolgen (siehe auch Abb. 5A).

B. Immunfluoreszenz von HeLa-Zellen, welche Luciferase und gp190^{S1} Tc-kontrolliert exprimieren.

In HtTA93-9-Zellen, welche die bidirektionale Transkriptions-einheit von (A) enthalten, wurde mit Antikörpern die Produktion von Luciferase (links), gp190^{S1} (Mitte) in Abwesenheit von Tc nachgewiesen. Nach Zugabe von Tc zeigte sich keine nennenswerte Synthese von gp190^{S1}, (wie dargestellt in Abb 5B, rechts).

C. Elektrophoretische Charkaterisierung von aus HeLa-Zellen aufgereinigtem gp190^{S1}.

HeLa-Zellklon HtTA93-9 sowie CHO-Zellklon CHO27-29 wurden mit oder ohne Tc kultiviert. Durch Elektrophorese aufgetrennte Zellextrakte wurden mittels "Western blot" mit mAK5.2 analysiert (Abb. 5C); links zeigt eine Analyse der CHO-, rechts der HeLa-Zelllinie. (1) = Kultur ohne, (2) = Kultur mit Tc, (3) = nicht transfizierte HtTA-1-Zelllinie, Molekulargewichtsstandards sind jeweils links angedeutet.

D. Aufreinigung von in Hela-Zellklon HtTA93-9 synthetisiertem gp190^{S1}

Die präparative Aufzucht der HtTA-93-9-Linie und Induktion der Expression von der gp190^{S1} durch Tc-Entzug erlaubte die Isolierung des Genproduktes über Affinitätschromatographie (mAK5.2-Säule).

Das Coomassie gefärbte Polyacrylamid-Gel (Abb. 6C) zeigte nach Elektrophorese ein Produkt, das aus gp190^{S1} sowie einem weiteren Protein von ca 50 kD bestand. Letzteres war kein Derivat von gp190^{S1}, stammte also aus HeLa-Zellen. Seine gezielte Abtrennung sollte jedoch keine prinzipielle Schwierigkeit darstellen.

Beispiel 4 Expression von gp190^S in Toxoplasma gondii und Aufreinigung des Produktes (siehe hierzu auch Abb. 6).

A. Die gp190^S-Sequenz wurde über MluI/PstI in den Vektor ppT eingesetzt. Damit wurde das Gen unter die Kontrolle des Tubulin-Promotors (Ptub-1) von *T. gondii* gebracht. Die 3' nicht-translatierte Region (VTR) stammte von dem Hauptoberflächenprotein von *T. gondii* (SAG-1) ab.

B. Expression von gp190^S in *T. gondii*.

Transfektion von *T. gondii* mit pTT190 führte zur Isolierung von Parasitenlinien, die konstitutiv gp190^S exprimierten. Die Im-

munfluoreszenz mit mAK5.2 (mittleres Bild) zeigte nicht nur die Expression des Gens, sondern legte auch die Verankerung des Expressionsproduktes an der Oberfläche des Parasiten nahe, da es, wie SAG-1, das gleiche Immunfluoreszenz-Muster erzeugte (rechter Teil der Abb. 6B); links in Abb. 6B ist eine Phasenkontrastaufnahme des mittleren Bildes dargestellt.

C. Isolierung von gp190^S aus T. gondii.

Aus präparativen Mengen von T. gondii (5×10^9 Parasiten) wurde gp190^S mittels Affinitätschromatographie (mAK5.2-Säule) aufgereinigt. Das hochreine Protein besaß das erwartete Molekulargewicht, wie das Coomassie-gefärbte Polyacrylamid-Gel nach Elektrophorese zeigte (2-3 aus Abb. 6C). Bei Nr. (1) Abb. 6C ist gereinigtes gp190^{S1} aus CHO-Zellen dargestellt mit Molekulargewichtsmarkierung an der linken Seite.

Beispiel 5: Charakterisierung des gp190^S mit monoklonalen Antikörpern.

Die Wechselwirkung von 16 monoklonalen Antikörpern mit gp190^S aus den verschiedenen heterologen Expressionssystemen wurde durch Immunfluoreszenz (IFA) an P. falciparum und T. gondii bzw. durch "Western blot" an den aufgereinigten Proteinen überprüft. Völlige Übereinstimmung wurde gefunden, wenn die beiden Parasiten verglichen wurden (Zahl der + deutet die relative Intensität der Fluoreszenz an). Im Western blot reagieren 12mAK's mit gp190^S aus E. coli und T. gondii. Im Gegensatz dazu binden 3 Antikörper nicht an das aus CHO-Zellen isolierte Material. Antikörper 15 und 16, die Epitope aus dem oligomorphen bzw. dem alternativen Allel (MAD20) erkennen, reagieren nicht. Die Ergebnisse sind in Tab. 1 zusammengefaßt, wobei ND = nicht durchgeführt bedeutet.

Beispiel 6:Expression des gp190^S in heterologen Systemen.1. Expression in E. coli

Das gp190^{S2} wurde in den Expressionsvektor, pDS56, RBSII eingesetzt, wo es unter Kontrolle des P_{N251aco-1}-Promotors stand, der über das lac Operator/Repressor/IPTG-System regulierbar ist (Abb. 4A). Die Überführung des Plasmids in Repressorproduzierende E. coli-Zellen, z.B. E. coli DH5 α Z1 erlaubt es, das gp190^{S2} unter IPTG-Kontrolle zu exprimieren. Das Produkt war aus dem Rohextrakt mittels einer Nickel-Chelatsäule über die durch den Vektor eingebrachte N-terminale (His)₆-Sequenz isolierbar. Eine anschließende Affinitätschromatographie an einer Antikörpersäule führte zu einem hochreinen Präparat. Da der verwendete monoklonale Antikörper (mAk5.2) ein konformationelles Epitop im C-terminalen Bereich erkannte, wird durch diese 2-Stufenreinigung auf intaktes Protein voller Länge, mit korrekter Faltung zumindest am C-Terminus, selektiert (Abb. 4B).

Das Endprodukt besitzt, im Gegensatz zu dem natürlichen Material, am N-Terminus 11 zusätzliche Aminosäuren, davon 6 Histidin. Es enthält keine N-terminale Signal- und auch keine C-terminale Anker-Sequenz. Die P. falciparum-spezifische Sequenz beginnt mit Aminosäure 20 und endet mit Aminosäure 1621.

2. Kontrollierte Expression des gp190^{S1} in HeLa- und CHO-Zellkulturen.

Das gp190^{S1} wurde in den Vektor pBi-5 eingesetzt und damit unter Kontrolle eines durch Tetrazyklin (Tc) regulierbaren Promotors gestellt. Das Tc-kontrollierte System wurde aus 2 Gründen gewählt:

- Es gehört zu den Expressionssystemen, mit denen höchste Ausbeuten in Säugerzellen erreicht werden.

- Nicht-sekretierte Fremdproteine in hoher Konzentration können mit dem Metabolismus der Zellen in negativer Weise interfe-rieren. Die Synthese des gewünschten Produktes wird daher erst nach Aufwachsen der Kultur initiiert.

In dem Konstrukt pBi5-gp190^{S1} wurde ein bidirektonaler Promotor durch Tc-kontrollierten Transkriptionsaktivator (tTA) akti-viert und intitierte die Transkription sowohl des gp190^{S1} als auch des Luziferase-Indikatorgens. In Anwesenheit von Tc ist der Promotor inaktiv. Die Transkriptionseinheit wurde sowohl in HeLa als auch in CHO-Zellen, welche beide konstitutiv tTA synthetisieren (HtTA-1-Linie (Gossen, M. and Bujard, H. (1992), Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. Proc. Natl. Acad. Sci. USA 89, 5547-5551); CHO-tTA-Linie, unveröffentlicht), überführt. Durch Kotransfektion (Ca²⁺-Phosphat-Methode) mit einem Hygromycin-Resistenz-vermittelnden Markergen wurde auf erfolgreiche chromosomalne Integration selektioniert. Hygromycin-resistente Klone wurden dann auf Regulierbarkeit der Expression \geq Tc un-tersucht, indem die Luziferaseaktivität als Indikator genutzt wurde. In gut regulierbaren Klonen (Regulationsfaktor \leq Tc 1000) wurde die gp190 Synthese geprüft. Immunfluoreszenz-Analyse (Abb. 5B) sowie Untersuchungen über "Western blot" (Abb. 5C) erlaubten, von beiden Zelltypen Klone zu identifizie-ren, welche gp190 unter streng regulierbaren Bedingungen syn-thetisieren. Von 20 Klonen wurde jeweils der bestregulierbare subkloniert. Die Subklone HtTA93-9 sowie CH027-29 wurden für Kulturen im 10 l Maßstab verwendet. Aus Zellextrakten dieser Kulturen ließ sich intaktes gp190^{S1} mittels Affinitätschromato-graphie (mAk5.2) isolieren. Das Material ist homogen bis auf eine einzige zelluläre Komponente, die nicht von gp190^{S1} ab-stammt und die etwa 25% des Präparats ausmacht (Abb. 6C). Sie müßte in einem weiteren Reinigungsschritt entfernt werden.

3. Expression des gp190^S in Toxoplasma gondii.

Toxoplasma gondii gehört wie *P. falciparum* zu den Apicomplexa und hat daher wahrscheinlich ein dem *P. falciparum* ähnliches Protein-Modifikationssystem. *T. gondii* lässt sich mit Fremd-DNA transfizieren, die effizient in das Genom integriert wird, außerdem lässt sich *T. gondii* problemlos in Zellkultur vermehren. Um ein möglichst natives gp190-Produkt zu erhalten, wurde gp190^{S2} so exprimiert, daß das Protein sekretiert und auf der Oberfläche des Parasiten, wie bei *P. falciparum*, über ein GPI-Motiv in der Membran verankert wird. Dazu wurde das gp190^{S2} (Abb. 3A) in das Plasmid ppTMCS (D. Soldati, unveröffentlicht) eingesetzt (Abb. 6A) und damit unter Kontrolle des *T. gondii*-Tubulin Promotors gestellt.

Dieses Expressionskonstrukt wurde in *T. gondii* transfiziert. Selektion mit Chloramphenicol führte zu resistenten Klonen, die gp190 synthetisieren, wie durch Immunfluoreszenz nachgewiesen wurde (Abb. 6B). Die Immunfluoreszenz mit anti-gp190-Antikörpern war nicht unterscheidbar von einer entsprechenden Anfärbung der Parasiten mittels Antikörper gegen SAG1, dem Hauptoberflächenprotein von *T. gondii*. Es ist daher davon auszugehen, daß gp190 an der Oberfläche von *T. gondii* verankert ist. Mehrere *T. gondii*-Klone (Nr. 3.1 bis 3.4) wurden charakterisiert und für die Produktion von gp190 aufbewahrt. Aus in präparativem Maßstab aufgewachsenen *T. gondii*-Kulturen (Klon 3.4) wurde gp190 mittels Affinitätschromatographie (mAK5.2.-Säule) isoliert. Analyse im elektrischen Feld zeigte ein homogenes Produkt mit einer Wanderungsgeschwindigkeit, die auf das intakte Protein schließen lässt (Abb. 6C).

Beispiel 7: Charakterisierung von gp190 Protein aus verschiedenen Expressionssystemen mittels monoklonaler Antikörper.

Ein Satz gp190-spezifischer monoklonaler Antikörper, von denen mehrere konformationelle Epitope erkennen, wurde dazu benutzt, über Immunfluoreszenz die Reaktivität der Antikörper mit *P.*

falciparum- bzw. T gondii-Parasiten zu vergleichen. Tabelle 1 zeigt, daß die Reaktivität der 16 Antikörper für beide Parasiten gleich ist. Dies ist ein starker Hinweis darauf, daß in T. gondii weitestgehend "natives" gp190 produziert wird. Der Vergleich der Reaktivität der Antikörper mit Protein aus E. coli, HeLa- bzw. CHO-Zellen sowie T. gondii zeigt ebenfalls, daß die meisten Antikörper mit den 4 Präparaten reagieren. Insbesondere wird das aus E. coli isolierte Protein von mehr Antikörpern erkannt als das Produkt aus Säugerzellen. Dies ist wahrscheinlich eine Konsequenz der Glykosylierung in Säugerzellen.

Beispiel 8: Immunsierung von Aotus lemurinus griseimembra-Affen mit gp190/MSP1 aus P. falciparum (FCB-1).

Zwei unabhängige Immunsierungsexperimente (A, B) wurden durchgeführt. Dazu wurde aus jeweils ca. 2×10^{11} Parasiten unter schonenden Bedingungen einmal 1,0 mg (A) und einmal 0,6 mg hochreines gp190/MSP1 isoliert.

Das Protein wurde zusammen mit Freund'schem Adjuvans (FCA) verabreicht. Die Kontrollgruppe erhielt lediglich FCA. Es wurde dreimal im Abstand von 4 Wochen mit gleicher Menge an Protein bzw. mit Adjuvans immunisiert. Zwei Wochen nach der letzten Immunisierung wurden die Tiere mit jeweils 10^5 Parasiten (FVO-Stamm) aus einem Donor-Tier infiziert. Die Parasitämie wurde täglich gemessen. Die Ergebnisse sind in Abb. 2 zusammengefaßt. Dabei bedeutet

T: daß die Tiere mit Resochin behandelt wurden

D: ein verstorbene Tier

Abb. 2A.: die Individuen der geimpften Gruppe erhielten je 3×60 Mikrogramm gp190/MSP1

Abb. 2B: die Individuen der geimpften Gruppe erhielten je 3×40 Mikrogramm gp190/MSP1

Während in den Kontrollgruppen nur 1/11 Tiere keine Parasitämie entwickelten, waren es in der geimpften Gruppe 6/10. Die vier Tiere aus der geimpften Gruppe, die eine hohe Parasitämie entwickelten, taten dies - im Vergleich zur Kontrollgruppe - mit einer durchschnittlichen Verzögerung von vier Tagen (Überschreiten der 2%-Grenze der Parasitämie).

Diese Experimente zeigten erstmals einen hochsignifikanten Schutz gegen Infektion mit *P. falciparum* durch gp190/MSP1 im Affenmodell. Das erfindungsgemäße Verfahren erlaubt es somit erstmals, einen wirksamen Impfstoff gegen die Malaria anzubieten.

Tabelle 1Wechselwirkung von gp190^S mit monoklonalen Antikörpern

Code	mAb	Art des Epitops	Variabilität	IFA				Western blot		
				P.f.	FCB	Toxoplasma	E. coli	Toxo- plasma	CHO	
1	5,2	konformationell	konserviert	+++	+++	+++	+	+	+	+
2	12,10	konformationell	konserviert	+++	+++	+++	+	+	+	+
3	7,5	konformationell	konserviert	+++	+++	+++	+	+	+	+
4	12,8	konformationell	konserviert	++	++	++	+	+	+	+
5	7,3	konformationell	dimorph (K1)	+++	+++	+++	+	+	+	+
6	2,2	konformationell	konserviert	+++	+++	+++	+	+	+	+
7	7,6	konformationell	dimorph (K1)	+++	+++	+++	+	+	+	+
8	9,8	konformationell	konserviert	+++	++	++	+	+	-	-
9	13,2	sequentiell	konserviert	+++	+++	+++	+	+	+	+
10	13,1	sequentiell	dimorph (K1)	+++	+++	+++	+	+	-	-
11	6,1	sequentiell	dimorph (K1)	+++	+++	+++	+	+	ND	ND
12	A5Z	nicht bekannt	nicht bekannt	+++	+++	+++	+	+	+	+
13	17,2	nicht bekannt	nicht bekannt	+++	+++	+++	ND	ND	ND	ND
14	15,2	nicht bekannt	nicht bekannt	+++	+++	+++	ND	ND	ND	ND
15	9,7	konformationell	dimorph (MAD20)	-	-	-	-	-	-	-
16	12,1	sequentiell	oligomorph	-	-	-	-	-	-	-

Patentansprüche

1. Herstellungsverfahren für das vollständige gp190/MSP1-Protein von *Plasmodium*, insbesondere *Plasmodium falciparum*, dadurch gekennzeichnet, daß das vollständige Gen für gp190/MSP1 in einem geeigneten System, vorzugsweise einem Wirtsorganismus, exprimiert wird.
2. Rekombinantes Herstellungsverfahren gemäß Anspruch 1, dadurch gekennzeichnet, daß der Synthese der dem Protein zugrunde liegenden DNA-Sequenz die DNA-Sequenz des *P. falciparum*-Stammes FCB-1 zugrunde gelegt wird.
3. Rekombinantes Herstellungsverfahren gemäß Anspruch 2, dadurch gekennzeichnet, daß der AT-Gehalt der exprimierten DNA-Sequenz gegenüber der natürlich vorkommenden Sequenz reduziert wurde, vorzugsweise von 74% auf 55%.
4. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz einschließlich Signalpeptid und Ankersignal kodiert.
5. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals kodiert.
6. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß das dem erzeugten Protein zugrunde liegende Gen für die volle Aminosäuresequenz mit Ausnahme des Ankersignals und des Signalpeptides kodiert.

7. Rekombinantes Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 6, dadurch gekennzeichnet, daß es folgende Schritte umfaßt:
 - (a) Entwurf der zu synthetisierenden DNA-Sequenz aus *P. falciparum* FCB-1, wobei eine DNA-Sequenz mit den im menschlichen Genom üblichen Codon-Häufigkeiten unter Erhalt der Aminosäuresequenz des FCB-1 Proteins hergestellt werden sollte,
 - (b) Einteilung der entworfenen Sequenz in überlappende Regionen, vorzugsweise in Regionen p83, p31, p36, gp30 und gp19,
 - (c) Synthese von Desoxyoligonukleotiden, die jeweils mindestens die gesamte Länge einer Region abdecken,
 - (d) Synthese der kodierenden Regionen für gp19, gp30, p36 und p31 durch PCR und Synthese der kodierenden Region für p83 durch Fusion aus zwei etwa 1200bp umfassenden Sequenzen,
 - (e) einzelne Klonierung der kodierenden Sequenzen
 - (f) Fusion des gesamten Gens und
 - (g) Expression in einem geeigneten Expressionssystem.
8. Rekombinantes Herstellungsverfahren nach Anspruch 7, dadurch gekennzeichnet, daß die in Schritt (c) synthetisierten Desoxyoligonukleotide durchschnittlich 120 Nukleotide lang sind und die benachbarten Sequenzen jeweils um ca 20 Basen überlappen.
9. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß als Expressionsvektor dPS56, RBSII verwendet wird.

10. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß als Expressionsvektor pBi-5 verwendet wird.
11. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 8, dadurch gekennzeichnet, daß als Expressionsvektor ppTMCS verwendet wird.
12. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß die dem Protein zugrunde liegende DNA-Sequenz in *E. coli* exprimiert wird.
13. Rekombinantes Herstellungsverfahren nach Anspruch 12, dadurch gekennzeichnet, daß der verwendete *E. coli*-Stamm der Repressor-produzierende Stamm *E. coli* DH5alphaZ1 ist.
14. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß die dem Protein zugrunde liegende DNA-Sequenz in HeLa-Zellen exprimiert wird.
15. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß die dem Protein zugrunde liegende DNA-Sequenz in CHO-Zellen exprimiert wird.
16. Rekombinantes Herstellungsverfahren nach einem oder mehreren der Ansprüche 1 bis 11, dadurch gekennzeichnet, daß die dem Protein zugrunde liegende DNA-Sequenz in *Toxoplasma gondii* oder *Leishmania* exprimiert wird.
17. Vollständige, zur Expression geeignete DNA-Sequenz des gp190/MSP1-Oberflächenproteins von *Plasmodium*, insbesondere *P. falciparum*, vorzugsweise erhältlich durch das rekombinante Herstellungsverfahren gemäß einem oder mehreren der Ansprüche 1 bis 16.

18. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 17, dadurch gekennzeichnet, daß sie nicht für das Ankersignal kodiert.
19. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 18, dadurch gekennzeichnet, daß sie weder für das Ankersignal noch für das Signalpeptid kodiert.
20. Zur Expression geeignete DNA-Sequenz gemäß Anspruch 19, dadurch gekennzeichnet, daß sie eine Sequenz für am N-Terminus vorliegende 11 zusätzliche Aminosäuren, davon 6 Histidine, umfaßt.
21. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 20, dadurch gekennzeichnet, daß die Sequenz keine erkennbaren "splice donor"- und "splice acceptor"-Signale enthält.
22. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 21, dadurch gekennzeichnet, daß die Sequenz keine größeren GC-reichen Sequenzen enthält.
23. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 22, dadurch gekennzeichnet, daß die Sequenz keine Erkennungssignale für Restriktionsenzyme, welche Sequenzen von sechs oder mehr Basenpaaren erkennen, enthält.
24. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 23, dadurch gekennzeichnet, daß die Sequenz für Erkennungssignale bestimmter Restriktionsnukleasen in Regionen, die die nach der Prozessierung des Proteins entstehenden Domänen trennen, einmal vorkommende Schnittstellen für Restriktionsendonukleasen enthält.
25. Zur Expression geeignete DNA-Sequenz gemäß einem oder mehreren der Ansprüche 17 bis 24, dadurch gekennzeichnet, daß

die Sequenz an ihren beiden Enden Schnittstellen für Restriktionsendonukleasen aufweist, die in der übrigen Sequenz und in einem zu verwendenden Vektor nicht vorkommen.

26. Wirtsorganismus, der die vollständige Nukleinsäuresequenz für das gp190/MSP1-Oberflächenprotein und/oder das vollständige Protein enthält.
27. Wirtsorganismus gemäß Anspruch 26, dadurch gekennzeichnet, daß der Wirtsorganismus *E. coli* ist.
28. Wirtsorganismus gemäß Anspruch 27, dadurch gekennzeichnet, daß der *E. coli*-Stamm der Repressor-produzierende *E. coli*-Stamm DH5alphaZ1 ist.
29. Wirtsorganismus gemäß Anspruch 26, dadurch gekennzeichnet, daß der Wirtsorganismus HeLa-Zellen sind.
30. Wirtsorganismus gemäß Anspruch 26, dadurch gekennzeichnet, daß der Wirtsorganismus CHO-Zellen sind.
31. Wirtsorganismus gemäß Anspruch 29 oder 30, dadurch gekennzeichnet, daß die Wirtszellen konstitutiv tTA synthetisieren.
32. Wirtsorganismus gemäß Anspruch 26, dadurch gekennzeichnet, daß der Wirtsorganismus *Toxoplasma gondii*, *Leishmania*, Baculoviren, Adenoviren oder Hefen vorgesehen sind.
33. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur aktiven Immunisierung gegen Malaria.
34. Verwendung eines gemäß einem der Ansprüche 1 bis 16 hergestellten gp190/MSP1-Proteins zur Herstellung von monoklonalen, zur passiven Immunisierung geeigneten Antikörpern.

35. Verwendung einer gemäß der Ansprüche 1 bis 16 hergestellten DNA-Sequenz zur Herstellung einer Vakzine auf Nukleinäurebasis.
36. Verfahren zur Stabilisierung von Gen-Sequenzen, dadurch gekennzeichnet, daß der AT-Gehalt der Sequenz verringert wird.
36. Stabilisiertes Gen, dadurch gekennzeichnet, daß es einen geringeren AT-Gehalt aufweist als das nicht stabilisierte Gen.
37. Vektor enthaltend eine DNA-Sequenz nach einem der Ansprüche 17 bis 25 und/oder 36.
38. Wirtszelle enthaltend einen Vektor nach Anspruch 37.
39. Impfstoff enthaltend ein Protein, hergestellt nach einem Verfahren nach einem der Ansprüche 1-16 und/oder eine DNA-Sequenz nach einem der Ansprüche 17-25 und/oder einen Wirt nach einem der Ansprüche 26-32 und/oder einen Vektor nach Anspruch 37.
40. Impfstoff nach Anspruch 39, dadurch gekennzeichnet, daß er weitere Immunität hervorrufende Produkte aus Plasmodium, insbesondere *P. falciparum*, enthält.

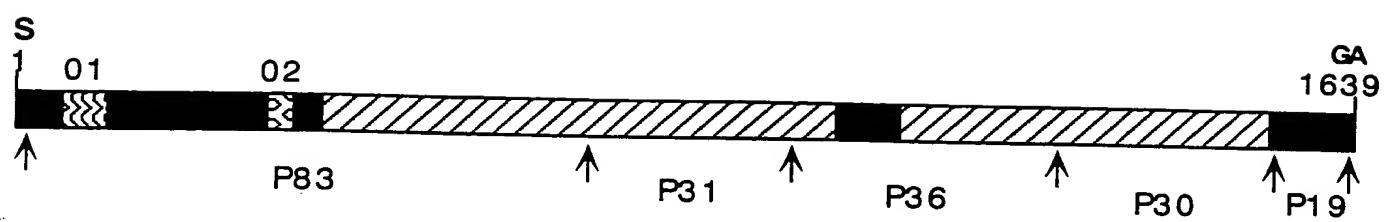
Zusammenfassung**Rekombinantes Herstellungsverfahren für ein vollständiges Malaria-Antigen gp190/MSP1**

Die Erfindung betrifft ein rekombinantes Herstellungsverfahren für das komplette gp190/MSP1-Oberflächenprotein von Plasmodium, insbesondere Plasmodium ~~falciparum~~, sowie die vollständige DNA-Sequenz dieses Proteins und geeignete Wirtsorganismen für die Expression der Sequenz, wodurch das Protein in seiner Gesamtheit erstmals außerhalb des Parasiten synthetisiert werden konnte.

Die Erfindung eröffnet erstmals die Möglichkeit, das gp190/MSP1-Oberflächenprotein in ausreichender Menge herzustellen; ferner ist es ein Gegenstand der Erfindung gp190/MSP1 als Impfstoff zur Verfügung zu stellen.

Schließlich gibt die Erfindung ein Verfahren zur Stabilisierung von AT-reichen Genen an (Abb. 1).

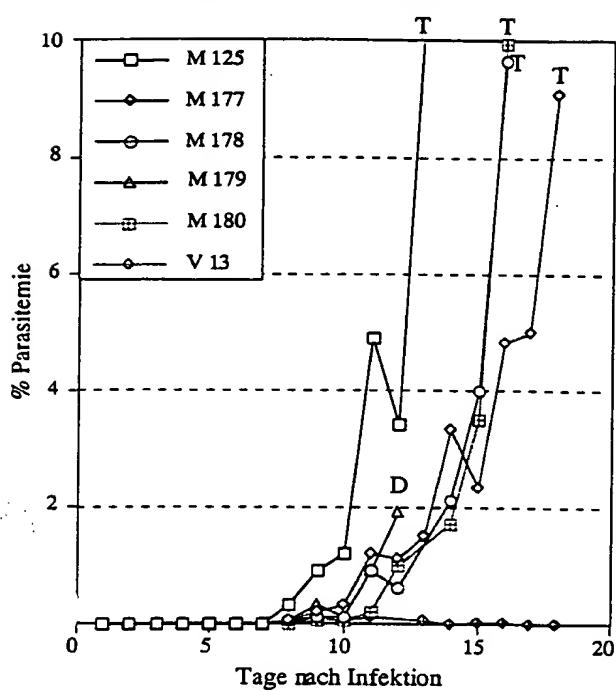
Abb. 1



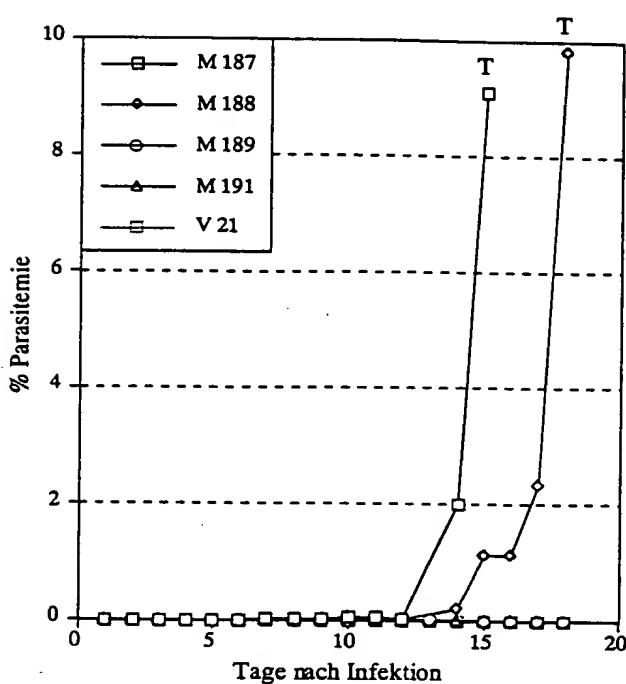
A

Abb. 2

KONTROLLGRUPPE

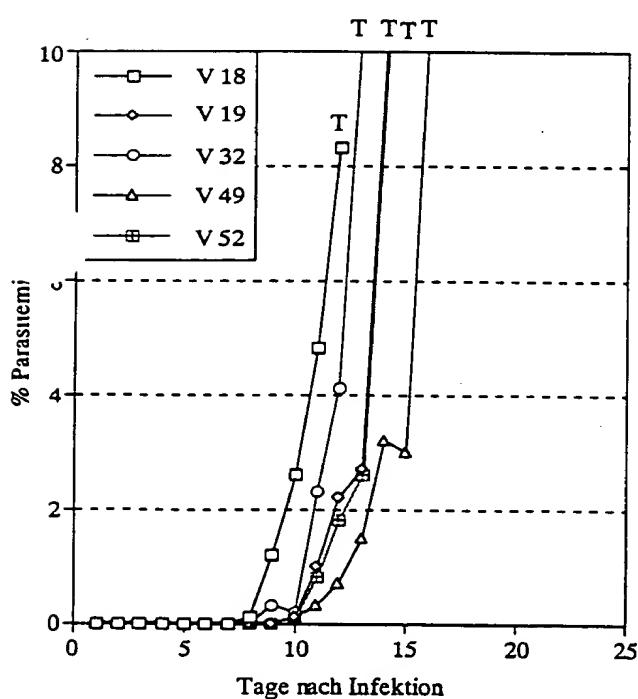


gp190 GRUPPE



B

KONTROLLGRUPPE



gp190 GRUPPE

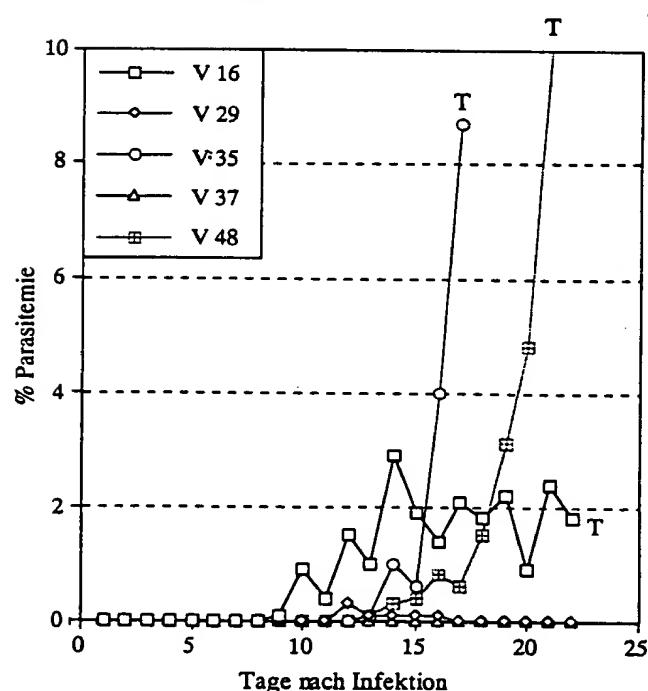


Abb.3A

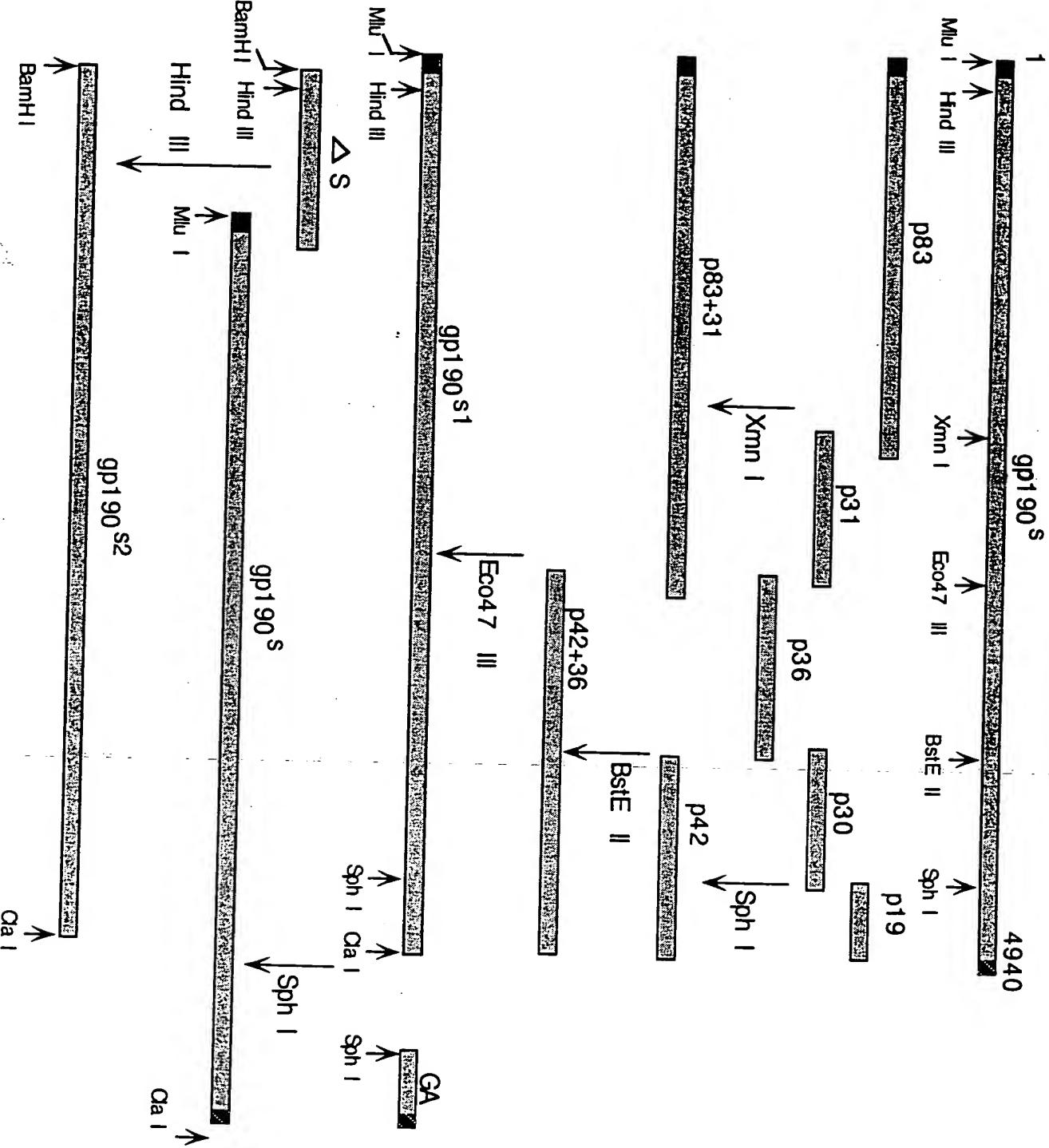
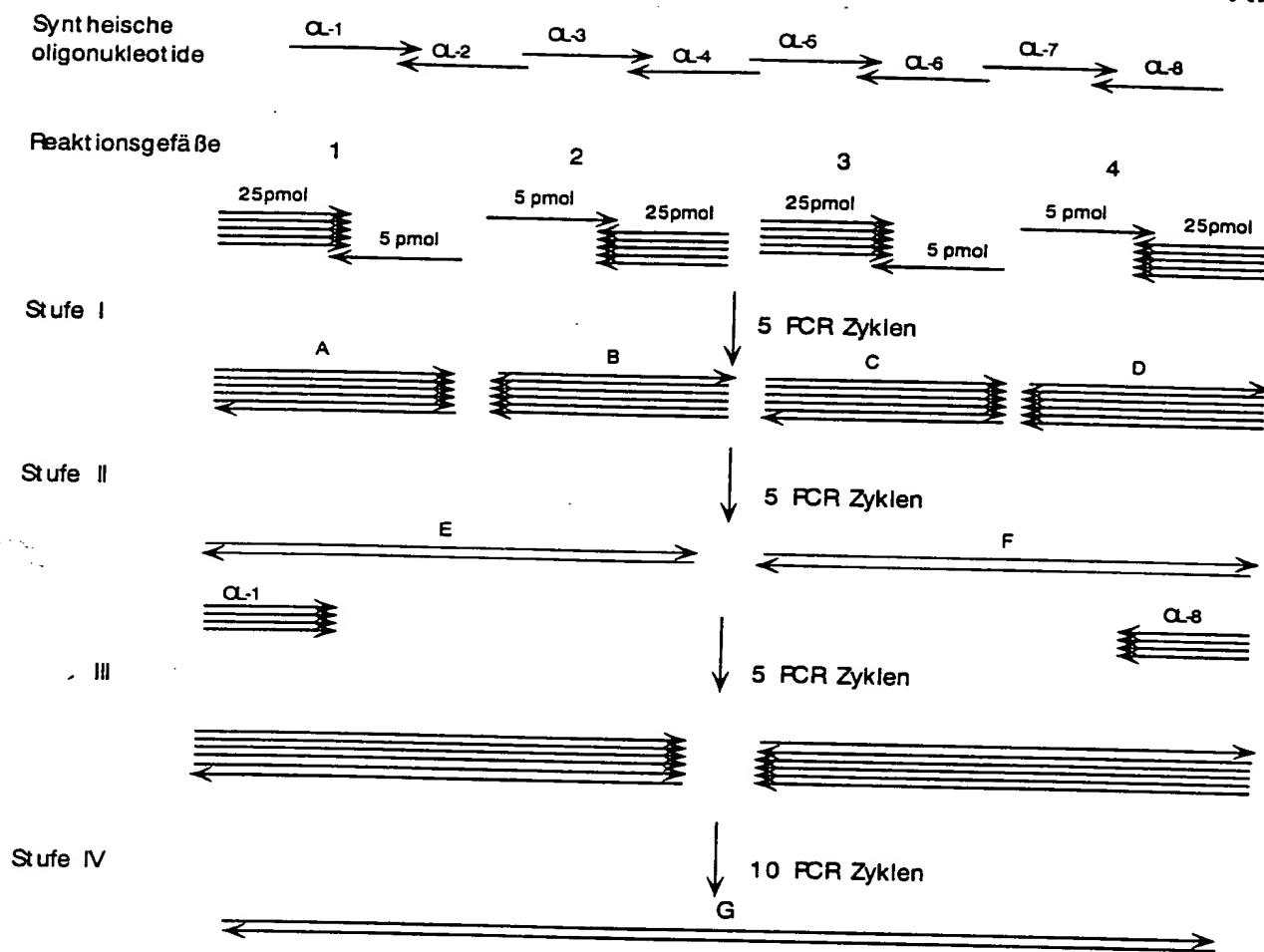


Abb. 3 B



**DNA Sequenz des nativen(gp190ⁿ) und des
synthetischen(gp190^s) Gens für gp190 aus FCB-1**

Abb. 3C

AS	M K I I F F L C S P L F F I I N T Q C V T H E S Y Q E	27
gp190 ⁿ	G A T T A T T A A A T A A T A G T A A A	
gp190 ^s	CGCACCGCTATGAAAATCATTTCTTCTCTGTCACTTCTGTTTATCATCAATACTCAGTCGCTGACCCACGAAATCCTATCAGAG MII	90
AS	L V K K L E A L E D A V L T G Y S L F Q K E K M V L N E G T	57
gp190 ⁿ	T C A A A A A T G A T T T T A T A A A A T A A A	
gp190 ^s	CTGGTTAAGAAACTGGAAGCTTGGAAAGATGCCCTTACCGGATAACGCCCTGTCAGAAGGAGAAGATGGTGTGAATGAAGGGACG	180
AS	S G T A V T T S T P G S K G S V A S G G G S G G S A S G G S	87
gp190 ⁿ	A A T T T T A G T A T T C A T A C A T T A T C A	
gp190 ^s	AGTGGCACGGCGTTACAACCAGCACCCGGTTCTAAAGGGCTGTGGCTAGGGTGGCTCCGGTGGCTGTGGCTCTGGGCTTC	270
AS	V A S G G S V A S G G S V A S G G G S G N S R R T N P S D N S	117
gp190 ⁿ	T T A T T C A T T T T T C A T A T T C A T A T T A	
gp190 ^s	GTGCCCTCCGGCGCAGCGTGGCATCAGTGGCTCAGTGGCAAGCGCGTCCCGGAAACAGTCGAAGAACCAATCCATCTGACAACTCT	360
AS	S D S D A K S Y A D L K H R V R N Y I L L T I K E L K Y P Q L	147
gp190 ⁿ	T A T T A T T T T A A A A C T C T G T A A A C A T T A C C	
gp190 ^s	AGCGATTCGACGCCAAGTCCTACGCCGACCTCAAGCACCAGTGAGAAACTATCTCCACTATCAAGGAGCTGAAGTACCCACAGTTG	450
AS	F D L T N H M L T L C D N I H G F K Y L I D G Y E E I N E L	177
gp190 ⁿ	T T T A T A T T T T A T A T A T A T A T A T A T A	
gp190 ^s	TTCGACCTCACTAATCATATGCTGACACTGTGTGATAACATTATGCTGATGACGGTTACGAAGAGATCAATGAACTC	540
AS	L Y K L N F Y F D L L R A K L N D V C A N D Y C Q I P F N L	207
gp190 ⁿ	T A T A A C T T T T A T A A A T A T A T T T T A T C T	
gp190 ^s	CTGTACAAGTTGAATTCTACTGACTTGTCAAGGGCAAACATGAATGACGTTGCGCAATGACTATTGTCAAAATTCCATTCAATTG	630
AS	K I R A N E L D V L K K L V F G Y R K P L D N I K D N V G K	237
gp190 ⁿ	A T C T A T A A C T A A C T G A A A A T A T T A T A A A	
gp190 ^s	AAAGATCAGGCCAACGAGTTGGACGTATTGAGAAGTTGGCTTCGGATATGCCAACCTCTCGACAAACATCAAGGACAATGGGAAAG	720
AS	M E D Y I K K N K K T I E N I N E L I E E S K K T I D K N K	267
gp190 ⁿ	C A A A A A A T A T A T A T A T A T A T A T A T	
gp190 ^s	ATGGAAGATTATATAAAAAGAATAAGAAGACCATCGAGAACATTAAACGAGCTGATCGAAGAACATCCAAAAGACCATAGACAAAATAAG	810
AS	N A T K E E E K K K L Y Q A Q Y D L S I Y N K Q L E E A H N	297
gp190 ⁿ	T A A A A A A A A T A T T T T T C T A T A A A T	
gp190 ^s	AATGCAACCAAGGAGGAAGAAAAGAAGAAGTTGTACCAAGGCCAGTACGACCTGTCCATCTATAACAAACAGCTGAGAACCCCCATAAC	900
AS	L I S V L E K R I D T L K K N E N I K E L L D K I N E I K N	327
gp190 ⁿ	T A A T T A A A T T T A A A C T G T A T T A A A A	
gp190 ^s	CTCATCAGCGTACTGGAGAACGCATAGACCCCTCAAGAACATGAAAAGACTGCTGACAAAGATTGAAATTGAAATTAAAGAAT	990
AS	P P P A N S G N T P N T L L D K N K K I E E H E K E I K E I	357
gp190 ⁿ	C A G T A T A A A T T C T T A A C A A A A A A A T	
gp190 ^s	CCTCCGCCAGCCAACCTCTGGAAACCCCTAACACGCTGCTGGACAAAGAACAGAACAGAGATAGGGAGCACGAGAACAGAGATCAAACAGATC	1080
AS	A K T I K F N I D S L F T D P L E L E Y Y L R E K N K N I D	387
gp190 ⁿ	T A T T T A G T A A A T A A T A A A A A T T	
gp190 ^s	GCCAAAACCATTAAGTCACATAGATTCTCTTACTGATCCCCCTGAGCTGAGACTACTTGAGAGAGAAGAATAAGAATATAGAC	1170
AS	I S A K V E T K E S T E P N E Y P N G V T Y P L S Y N D I N	417
gp190 ⁿ	A A G T A G T A T C A A T T T T A A T A A T T T	
gp190 ^s	ATCTCCGCCAAAGTCGAGAACAGAACCTAATGAATATCCAATGGTGTGACCTGACCTCTGTCTTATAACGATATCAAC	1260
AS	N A L N E L N S F G D L I N P F D Y T K E P S K N I Y T D N	447
gp190 ⁿ	T T A T A T T C T T A A T A T A A A A A G A C A T T T	
gp190 ^s	AACGCTCTCAACGAGCTAACATTCTCGTACTTGAATTAAACCCCTCGATTATACGAAAGAACCCCTAACAGAACAT	1350
AS	E R K K F I N E I K E K I K I E K K K I E S D K K S Y E D R	477
gp190 ⁿ	A A A C A T T A A T A A A A A T C T A T C T A T A T A T	
gp190 ^s	GAGAGAAAGAAGTTATCAACGAAATCAAGGAGAACATCAAATTGAGAGAGAAGAAAATTGAGAGTGACAAAGAAAAGTTACGAAGACCC	1440
AS	S K S L N D I T K E Y E K L L N E I Y D S K F N N N I D L T	507
gp190 ⁿ	T C T G T C T T A A A A T A T T A T A G T A T T A T A T	
gp190 ^s	AGCAAAAGTCTAACCGATATCACTAAAGAGTATGAAAAGCTGCTGACGGAGATCTATGATCCAAATTCAACATAACATCGACCTGACC	1530
AS	N F E K M M G K R Y S Y K V E K L T H E N T F A S Y E N S K	537
gp190 ⁿ	T A T A T A T T T T T T T A A A A A A	
gp190 ^s	AACTTCGAGAAAATGATGGGAAACGGTACTCTTACAAAGTGGAGAAACTGACACACCATAATACCTTGCATCTATGAGAATTCTAAG	1620

AS	H N L E K L T K A L K Y M E D Y S L R N I V V E K E L K Y Y	567
gp190 ⁿ	A T A A A A T T A A T A A T A A T A T	
gp190 ^s	CATAATCTTGAGAAGCTCACCAAGCTTAAAGTATATGGAGGACTATTCTCGGGAACATTGGTGTGGAGAAAGAACTAAAGTATTAC	1710
AS	K N L I S K I E N E I E T L V E N I K K D E E Q L F E K K I	597
gp190 ⁿ	A T A C A A T T A A T A A T A A C T A A A	
gp190 ^s	AAGAAATCTCATAAGTAAGATCGAAAACGAGATCGAGACGGCTTGTGAGAACATTAAGAAGGATGAAGAACAGTTGTGAGAACAGATT	1800
AS	T K D E N K P D E K I L E V S D I V K V Q V Q K V L L M N K	627
gp190 ⁿ	T A A T A A A T C A A A T A A T A A T A T A T A A	
gp190 ^s	ACAAAAGACGAAAATAACCGAGATGAGAACAGATCTGGAGGTCTCGATTTGTTAAAGTCCAAGTGCAGAAGGTGCTCTCATGAACAAG	1890
AS	I D E L K K T Q L I L K N V E L N H N I H V P N S Y K Q E N	657
gp190 ⁿ	C T A A T G T A A T A A T C T C T C C A A A	
gp190 ^s	ATTGATGAACTCAAGAACGACTCAACTCATCTGAAGAACGTGGAGTTAAACATAATATACATGTGCCAAGTAGTTAACCGAGGAAAT	1980
AS	K Q E P Y Y L I V L K K E I D N L K V F M P K V E S L I N E	687
gp190 ⁿ	A T T T A T T G T G A A T T T A T G A A T C A T A T	
gp190 ^s	AACCAAGAACCATACTACCTCATCGTACTCAAGAACAGAGATAGACAAACTGAAAGATGTTCATGCCAAGTGGAGACCTGATCACCGAA	2070
AS	E K K N I K T E G Q S D N S E P S T E G E I T G Q A T T K P	717
gp190 ⁿ	A A A A A T A T A G T G A A A A C A A A A A T A A T	
gp190 ^s	GAGAAGAAGAACATTAACCTGAAGGACAGTCAGATACTCCGAGCCTTCCACAGAAGGAGAGATAACCGGACAGGCTACCCACAGCCC	2160
AS	G Q Q A G S A L E G D S V Q A Q A Q E Q K Q A Q P P V P V P	747
gp190 ⁿ	A A A T T A A A T C A A A A A A A A A A A A A A A A	
gp190 ^s	GGACAAACAGGCCGGTTCAGCTCTCGAACCGATAGCGTGCAGCTCAAGCACAAAGAGCAGAACAGCAGCAGCCAGCTCCAGTGCCAGTGCCC	2250
AS	V P E A K A Q V P T P P A P V N N K T E N V S K L D Y L E K	777
gp190 ⁿ	A A A A A C A A A A T A T A T T C C T A T T A A A A A A A A	
gp190 ^s	GTTCCAGAGGCTAAAGCTCAAGTGCCTACACCACAGCTCTGTGAATACAGAACCGAGAATGTCAGCAAATGGACTACCTTGAGAAAG	2340
AS	L Y E F L N T S Y I C H K Y I L L V S H S T M N E K I L K Q Y	807
gp190 ⁿ	T A A T T A T A T A T T G T A T C A A A A A T A A T A A T	
gp190 ^s	CTCTATGAGTCTCTGAATACATCTCACATCTGCCACAAATATACATCTCTCTCACAGCACTATGAAACGAGAAGATTCTAAACAGTAC	2430
AS	K I T K E E E S K L L S S C D P L D L L F N I Q N N I P V M Y	837
gp190 ⁿ	A T A G A A C T A A G T A A T A T A T A A T A A T A A T A	
gp190 ^s	AGATAACCAAGGAGAGGAGTAAACTGTCCTTGTATGTCATGGACCTGCTGTCAATATCCAGAACACATTCCGTTATGTAT	2520
AS	S M F D S L N N S L S Q L F M E I Y E K E M V C N L Y K L K	867
gp190 ⁿ	T T A A G T A A A T A T A T A A A T T T A T T A T G	
gp190 ^s	TCTATGTCGATAGCCTCAACATTCTCTCTCAACTGTCATGGAGATATGAGAAGGAGATGGTCTGCAACCTGTATAAACTCAAA	2610
AS	D N D K I K N L L E E A K K V S T S V K T L S S S S M Q P L	897
gp190 ⁿ	T T A A A T T A T A G A A A A A A T A A G T T C A A T A	
gp190 ^s	GACAAAGCACAAAGATTAAGAACCTCTGGAGGAAGCTAAGAACGGTCTCCACCTCTGTAAACTCTCTCCAGCTCATGCAACCACTG	2700
AS	S L T P Q D K P E V S A N D D T S H S T N L N N S L K L F E	927
gp190 ⁿ	A T A G T A A T A T T T A A T T A T T G T A G T T A T A A	
gp190 ^s	TCTCTCACACCTCAAGACAACCCCGAAGTGAGCCTAACGAGCACCTCTCACTGACCAACCTTAATAACTCAACTGAAACTGTTGAG	2790
AS	N I L S L G K N K N I Y Q E L I G Q K S S E N F Y E K I L K	957
gp190 ⁿ	A T A G T A A C A T A T A A T A A T A A G T A G T A T T A T A	
gp190 ^s	AAACATCTGTCTCGCGCAAGAATAAGAACATCTACCAAGAACATTGGACAGAAATCTGCGAGAACCTTACAGGAGAAGATCTGAA	2880
AS	D S D T F Y N E S F T N F V K S K A D D I N S L N D E S K R	987
gp190 ⁿ	T T T T A T T A T A T T A T A T T A T T A T G T A A A G	
gp190 ^s	GACAGCGACACATTCTATAACGAGAGCTCAACTCTCTGTAAAGCCGATGATATCAACTCTCTAACGATGAAATCTAAACGATGAAATCTAAACG	2970
AS	K K L E E D I N K L K K T L Q L S F D L Y N K Y K L K L E R	1017
gp190 ⁿ	A T A A T T A T A A A A T T A G T A T C A T T A T T A T A T A A T A	
gp190 ^s	AAGAACGCTGGAAGAGGACATCAATAAGCTGAAGAACACACTGCAACTGACCTCTGCACCTGTACAACAGTACAAACTGAAACTGGAGAGA	3060
AS	L F D K K R T V G K Y K M Q I K K L T L L K E Q L E S K L N	1047
gp190 ⁿ	T A T T A A T T A A A T A A A C T T A T A A A A T A T C A T G T	
gp190 ^s	CTCTTCGACAAGAAGAACAGCTGGCAAGTATAAGATGCGAGATCAAGAACGTTGACTCTGCTCAAGGAGCAGCTGTGAAAGCAAACTCAAC	3150
AS	S L N N P K H V L Q N F S V F F N K K K E A E I A E T E N T	1077
gp190 ⁿ	T T C A G T T A A T T T T A A A T A A A T A A A T A A A A	
gp190 ^s	TCACTGAACAATCCGAAACACGTACTCCAGAACCTCTCAACTGTTCTCAACAAAGAACAGGAGCCGAGATCGCCGAGACAGAGAACACT	3240
AS	L E N T K I L L K R B Y K G L V K Y Y N G E S S P L K T L S E	1107
gp190 ⁿ	T A A A A A T A T G T T A T T A T A A T A A T A A T A A T A	
gp190 ^s	CTGGAGAACACCAAGATTCTCTCAACACTACAAAGGCCCTGTCAGTATTATAATGGCAGATCTCTCTCTGAAGACTCTCTCCGAG	3330
AS	E S I Q T E D N Y A S L E N F K V L S K L E G K L K D N L N	1137
gp190 ⁿ	ATCA T A A A T T T A A T A A T A A T A A T A A T T A T	
gp190 ^s	GAGAGCATCCAGACCGAGGATAACTACGCCAGCCCTGAGAACCTCAAGGTCTGTCTAAGCTCGAAGGCAAGCTGAAGGACAACCTGAC	3420

AS	L E K K K L S Y L S S G L H H L I A E L K E V I K N K N Y T	1167
gp190 ⁿ	T A A A A A T A T C A T A A T T A T T A T A A A A A T A T T A	
gp190 ^s	CTGGAGAAGAAGCTCACCTCTCTACCGGACTGCATCACCTGATGCCAGCTCAAGGAAGTCATTAAGAACAGAACTACACC	3510
AS	G N S P S E N N T D V N N A L E S Y K K F L P E G T D V A T	1197
gp190 ⁿ	T T C T T A G T T C T T A A A T C A T A A T A A T T A A	
gp190 ^s	GGCAATAGCCCACCGAGAAATAACAGACGTGAATAACCGCACTGGAACTCTACAAAGAAGTCTCTGCCTGAGGAACAGATGTCGCCACT	3600
AS	V V S E S S G S D T L E Q S Q P K K P A S T H V G A E S N T I	1227
gp190 ⁿ	T A A G A G A T A A A A A A A A A A A A A A A T C A	
gp190 ^s	GTGGTGTCTGAATCTGGCTCCGACACACTGGAGCAGTCACCTAACCTAACGGCTGCATCTACTCATGTCGGAGCCGAGTCCAAATCAATT	3690
AS	T T S Q N V D D E V D D V I I V P I P G E S E E D Y D D L G	1257
gp190 ⁿ	A A A A T T A A A A A A A A A A A A A A A T C A	
gp190 ^s	ACCACATCTCAGAACGTCGACGATGGTCGATGACGTACATCTGTCCTATCTCGCGAGGCGAGGAGACTACGATGACCTCGGC	3780
AS	Q V V T G E A V T P S V I D N I L S K I E N E Y E V L Y L K	1287
gp190 ⁿ	A A A A A A A A A A A A A A A A A T T T A T T T G T T A T A	
gp190 ^s	CAGGTGGTCAACCGTGAAGGCTGTCACTCCCTCCGTGATTGATAACATTCTGTCAAATCGAGAACGAAATCGAACGATGTCCTATCTGAAA	3870
AS	P L A G V Y R S L K K Q L E N N V M T F N V N V K D I L N S	1317
gp190 ⁿ	T A T T A A A G T A A A T A A A T A T T T T T T A T C A	
gp190 ^s	CCCTCGCAGGGCTCTATAGCTCTCAAGAACAGCTGGAGACCTGATTCCCTATAAGACCTGACCTCTAACTACGTTGTCAAG	3960
AS	R F N K R E N P K N V L E S D L I P Y K D L T S S S N Y V V K	1347
gp190 ⁿ	A A C T A T T A A T C A T A A A T T A A A A A A A A A T T A	
gp190 ^s	CCCTTTAATAAGAGAGAAAATTCAAGAACGCTGGAGAGGACTGATTCCCTATAAGACCTGACCTCTAACTACGTTGTCAAG	4050
AS	D P Y K F L N K E K R D K F L S S Y N Y I K D S I D T D I N	1377
gp190 ⁿ	T T A T T A A A A C T A A G C T T T T A A T A A T G A	
gp190 ^s	GACCCATACAAGTCTCTCAATAAAAGAGAACAGGGATAAATTCTGTCTAGTTACAACATATCAAGGACTCCATCGACACCGATAATCAAT	4140
AS	F A N D V L G Y Y K I L S E K Y K S D L D S I K K Y I N D K	1407
gp190 ⁿ	T A T T A T A A T A T C A T A A T T A T A A A C A	
gp190 ^s	TTCGCTAATGATGTCGCTGGGTATTACAAGATCTGAGCAGAAAATACAAGCTGACCTTGACTCTATTAAAAAGTATATCACACGATAAG	4230
AS	Q G E N E K Y L P F L N N I E T L Y K T V N D K I D L F V I	1437
gp190 ⁿ	T A G C T T T A C T T G T A T A T T T T T T T A T T A T	
gp190 ^s	CAAGGGCAGAACGAAAATATCTGCCCTCCGTGAATAACATCGAAACCCGTACAAGAACAGTGAACGACGAAAATCGACCTCTCGTAATT	4320
AS	H L E A K V L N Y T Y E K S N V E V K I K E L N Y L K T I Q	1467
gp190 ⁿ	TT A A A A T A T A T A T C A C A A A A T T T A T T A T	
gp190 ^s	CACCTGGAGGCCAGGTCTCAACTATACGAGAACGAAATGAGGAACTACGAGCTGAACACTCTCAAAACATCCAA	4410
AS	D K L A D F K K N N N F V G I A D L S T D Y N H N N L L T K	1497
gp190 ⁿ	A T T A T A T T T T T A A A A T T C T A T A T T C T A T A	
gp190 ^s	GACAAGCTGGCAGATTCAAGAAAAATAACAAATTCTGTCGGATTGAGACCTGTCACCGATTATAACCCACAATCTCTGACCAAG	4500
AS	F L S T G M V F E N L A K T V L S N L L D G N L Q G M L N I	1527
gp190 ⁿ	C T A G T A T T T T C T T A T C T T A T T A T A T T A T	
gp190 ^s	TTCTGTCCACTGGCATGGTGTGCAAAACCTGCCAACAGTGTCTGAGCAATCTGTCAGCGAACCTGCAAGGCATGCTGACATC	4590
AS	S Q H Q C V K K Q C P Q N S G C F R H L D E R E E C R C L L	1557
gp190 ⁿ	A A A A A A A T A A A T C T A A A T A T A A A T A T A T A	
gp190 ^s	TCCCAGCACCAGCGTGAAGAACAGTGGCCCGAACAGTGTCTGAGCAATCTGTCAGCGAACCTGCAAGGCATGCTGACATC	4680
AS	N Y K Q E G D K C V E N P N P T C N E N N G G C D A D A K C	1587
gp190 ⁿ	T T A T T A T T T T C T A T A T A T A C T A T A C T	
gp190 ^s	AACTACAAACAGAACAGAGATACTGCGTGAAGAACCCAAACCTACCTGCAATGAAAACATGGGGGTGTGAGCCGATGCTAAATGC	4770
AS	T E E D S G S N G K R I T C E C T K P D S Y P L F D G I F C	1617
gp190 ⁿ	A T T C A T A G C T A T A T T T T T T T C	
gp190 ^s	ACCGAGGAAGAACAGCGCTCTAACGGAAAGAACATCGGAGTGTACTAACGCCGACTCTATCCACTCTCGACGGGATTTC	4860
AS	S S N F L G I F F L L I L M L I L Y S F I * *	1639
gp190 ⁿ	AGCTTC C T A A C A T A T A A T A A T A A T T T T T C	
gp190 ^s	TCCAGCTCTAATTCTCTGGCATCTCTGTCATGCTGATCTCATGCTGATCTGACAGCTCTATCCACTCTCGACGGGATTTC	4940
	stop codon C1a 1	

Abb.3D

N'-terminus

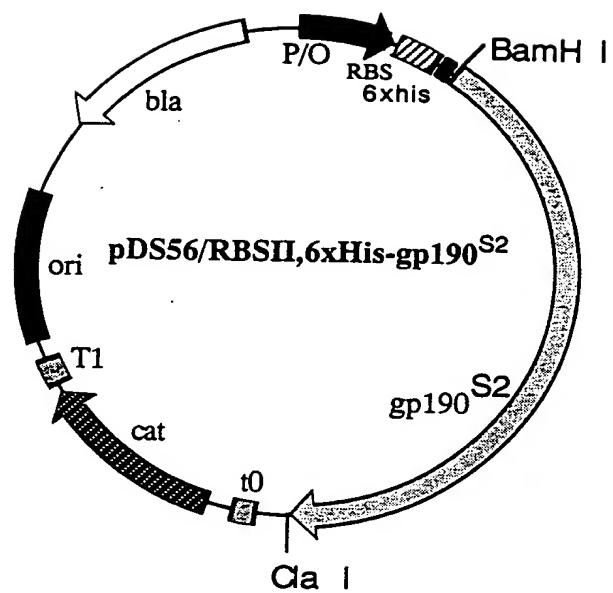
C'-terminus

gp190s1 Sequence

DNA Sequence	GC	ACCGCGTATGAAATC	-----	AGCTCTTAATAAGGGGGCATCGATGGC
AA Sequence	Mlu I	Met Lys Ile	-----	Ser Ser Asn stop codon
AA Position	1	2	3	Not I
				Cla I
				1619 1620 1621

gp190s2 Sequence

DNA Sequence	GC	GGATCCCGTGACCCAC	-----	AGCTCTTAATAAGGGGGCATCGATGGC
AA Sequence	BamH I	Val Thr His	-----	Ser Ser Asn stop codon
AA Position	20	21	22	Not I
				Cla I
				1619 1620 1621



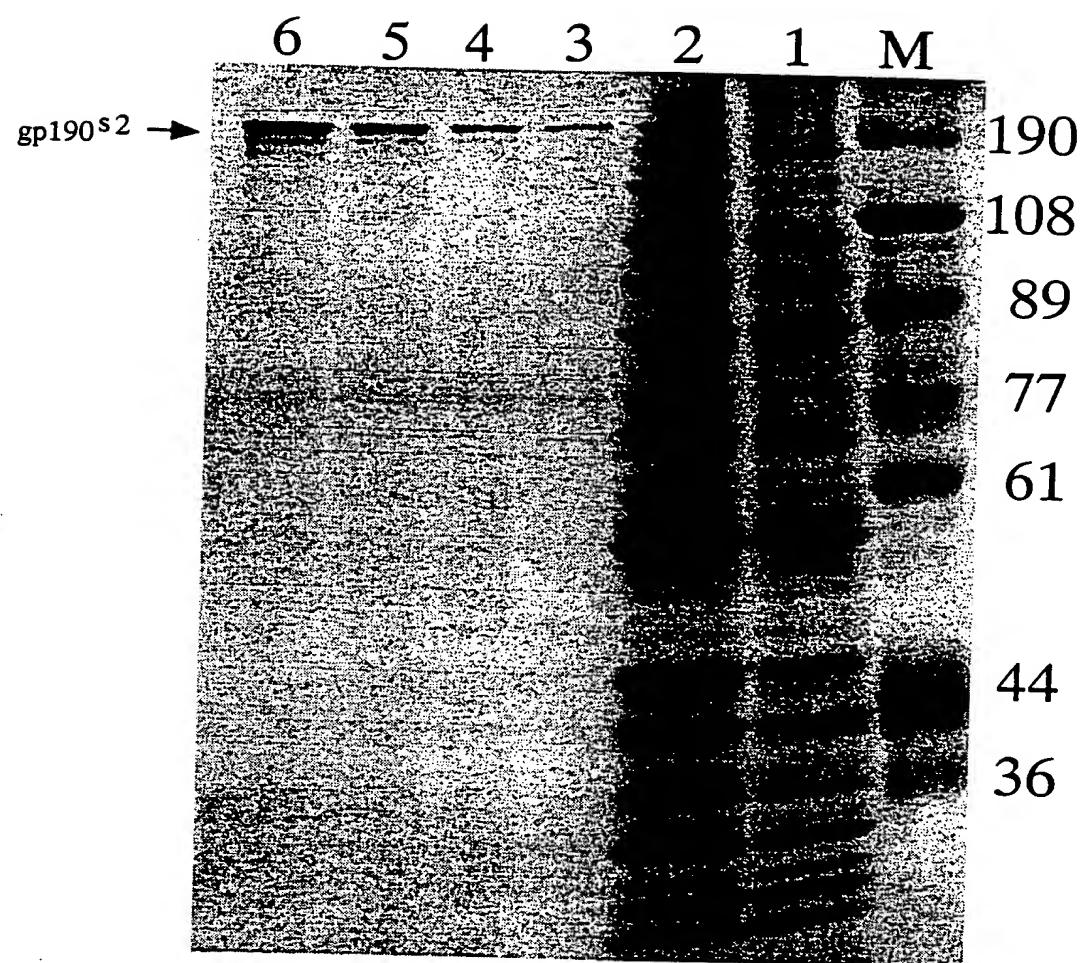


Abb. 5A

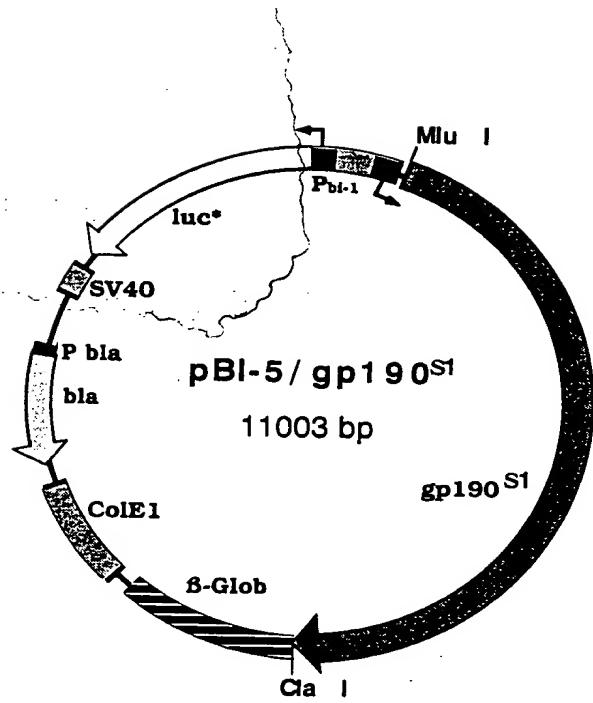


Fig. 5B

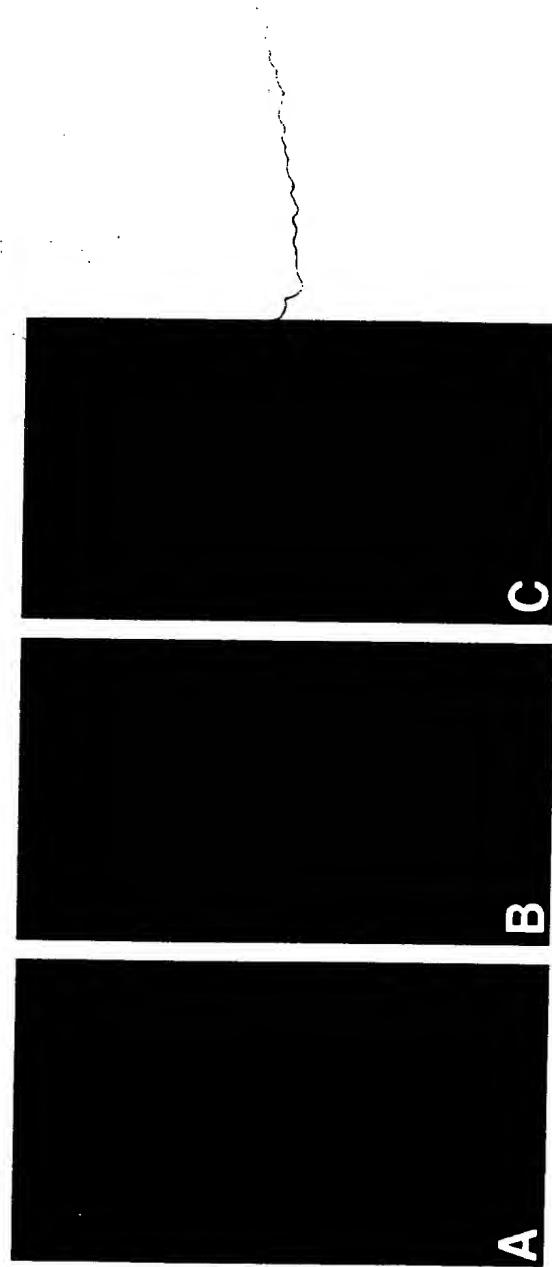


Abb.5C

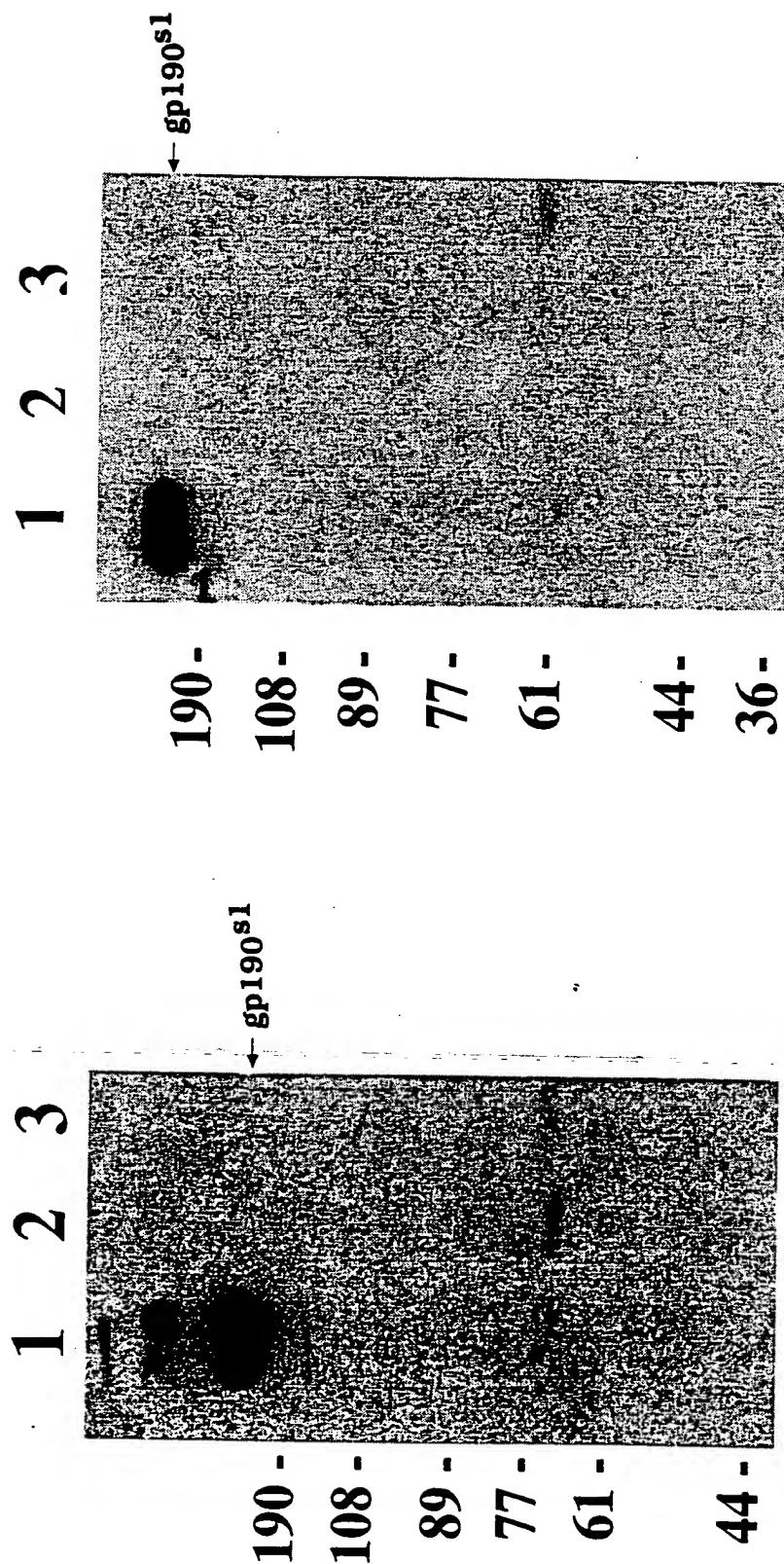


Abb. 6A

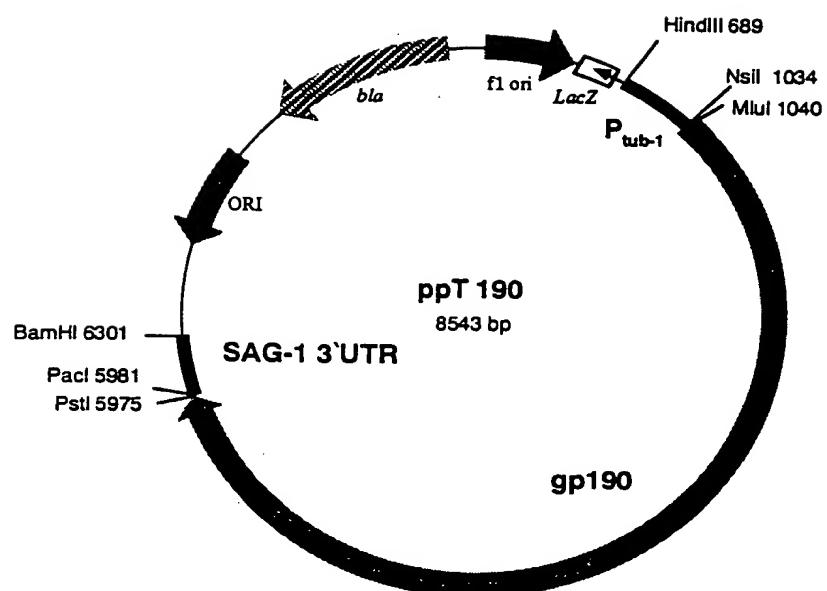


Fig. 6B

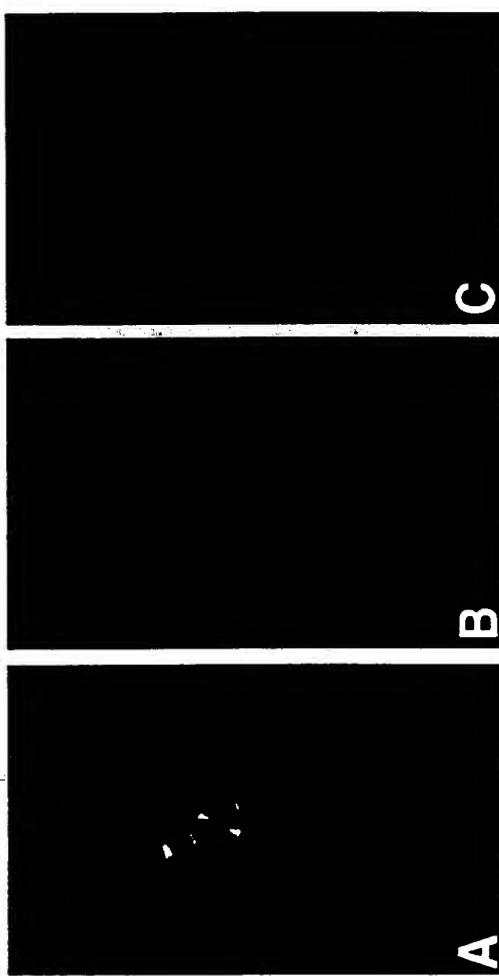
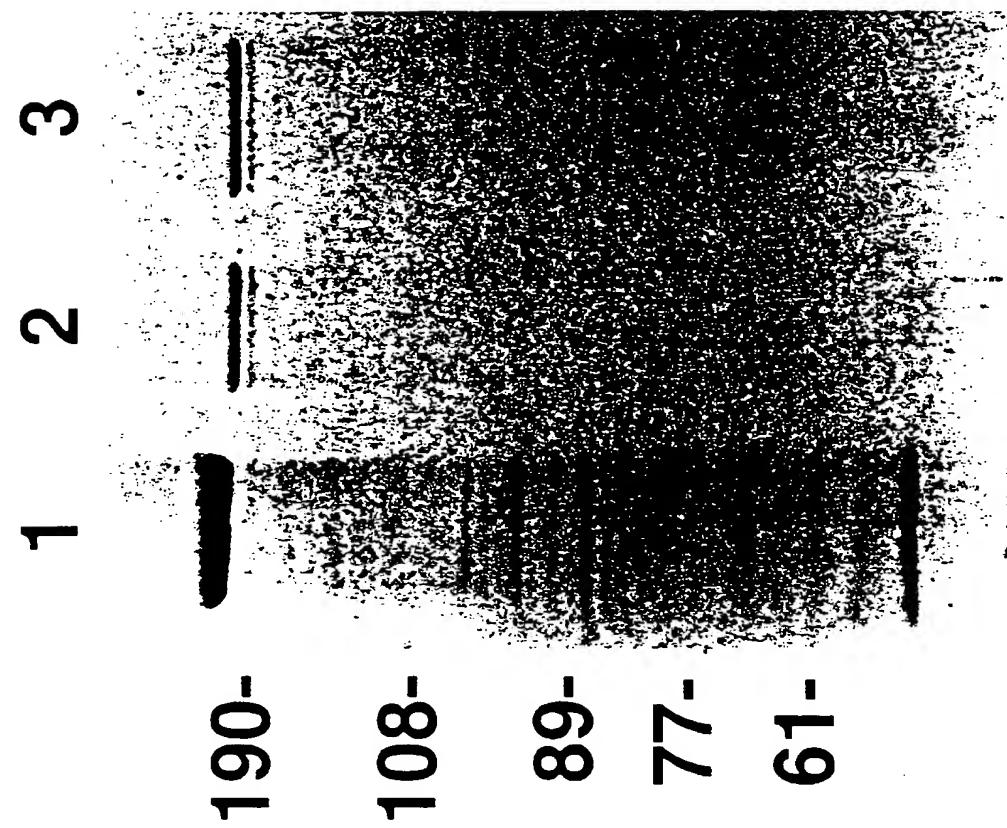


Fig. 6C



THIS PAGE BLANK (USPTO)